Introductory notes to General Relativity

Manuel Tiglio
Center for Scientific Computation and Mathematical Modeling,
Department of Physics,
Joint Space Sciences Institute.
Maryland Center for Fundamental Physics,
University of Maryland
College Park, MD 20742, USA
email: `tiglio_AT_umd.edu`

Abstract

Notes for four lectures on an introduction to GR at the *NITheP 24th Chris Engelbrecht Summer School 2013: Gravitational Wave Astronomy*. Rhodes University, South Africa, January 15-24 2013. The notes draw from Wald's and Carroll's books, with some problems from them and from the *Problem Book in Relativity* and Gravitation by Lightman, Press, Price and Teukolsky, and from *Liv. Rev. in Relativity 15 (9) 2012* by Sarbach and Tiglio.

Contents

1 Introduction

In General Relativity (GR) events take place in spacetime. That is, an event is both a place and an instant of time where something happens. For the moment being you can think of spacetime as that one of Special Relativy, $\mathcal{R}^4 = \mathcal{R} \times \mathcal{R}^3$ ("time $\times$ space"). In GR spacetime is a four-dimensional manifold. For all practical purposes we can think of an n-dimensional manifold as defined by the property that locally (but not necessarily globally) can be described by coordinates in $\mathcal{R}^n$ (or mapped to a region of $\mathcal{R}^n$). If you want to think of examples of manifolds which are not linear spaces a sphere or a torus (both of which are 2-dimensional manifolds) would do it.

In the General Relativity case the dimensionality of the manifold will be $n = 4$ but in much of the differential geometry description we will keep such dimension arbitrary since you might find it useful in other circumstances (for example, Newtonian physics in curvilinear coordinates).

In Newtonian physics and even in Special Relativity we are used to identifying vectors with points in space and spacetime, respectively. For example, in Newtonian physics there is an absolute concept of preferred time and events take place in space, which we usually take to be $\mathcal{R}^3$. A point in space needs three coordinates to be described and so does a vector, therefore we tend to think of them as the same kind of mathematical objects. Something similar happens in Special Relativity, depending on the point of view of the person. In GR we do need to go beyond this notion and distinguish between the spacetime (which in general will not be a vector space but a manifold) and how to do calculus on it. The latter will lead to the definition of the tangent space at each point in the manifold, parallel transport and, eventually, curvature and the Einstein equations.

We start with some notation which will we use here (and which is pretty standard also), to which you can refer once they are more than names (i.e. have been defined).

2 Notation and conventions

- GR=General Relativity.
- $\mathcal{R}$: the set of real numbers
- $\mathcal{M}$: the spacetime, which will be a four-dimensional manifold. Roughly speaking, an $n-$dimensional manifold can be defined as a space (in general *not* linear) which can be locally mapped to $\mathcal{R}^n$ (for example, the sphere S^2 cannot be mapped to $\mathcal{R}^2$ by a single map - this is a consequence of being a compact space- but it can be locally mapped to $\mathcal{R}^2$) . In the case of special relativity $\mathcal{M} = \mathcal{R}^4$.
- $\mathcal{F}$: functions on the manifold. That is, $f \in \mathcal{F}$ is a map

$$f : \mathcal{M} \to \mathcal{R} .$$

- $\{x^\mu\}$: some set of coordinates for the manifold or a part of it.
- T_p: the tangent space at a point $p \in \mathcal{M}$. It is a linear space with the same dimensionality of the manifold. It will be defined as the set of directional derivatives of functions on $\mathcal{M}$, see Section 4.1. That is, if $v \in T_p$ and $f \in \mathcal{F}$ then

$$v(f) := \left(\sum_{\mu=1}^{n} v^\mu \frac{\partial}{\partial x^\mu} f \right) \Bigg|_p \tag{1}$$

  and $\{v^\mu \, , \mu = 1, \dots, n\}$ (notice the upper indices) are the *components* of the vector v on the *coordinate basis* of T_p induced by the coordinates x^μ.

- The repeated indices convention: when an upper and a lower index are repeated (have the same letter) a summation is implicitly assumed and the $\sum$ symbol usually skipped. For example, the above equation would read

$$v(f) = \left(v^\mu \frac{\partial}{\partial x^\mu} f \right) \Bigg|_p .$$

  If there are repeated indices at the same "level" (both up or both down) there is something wrong.

- Furthermore, the following abbreviation is many times used:

$$\partial_\mu := \frac{\partial}{\partial x^\mu} ,$$

  with which Eq. (1) would just read $v(f) = v^\mu \partial_\mu f$ (where we have skipped, and will often do so, the explicit evaluation at p because in many cases v will be a *vector field*, defined on –at least– a neighborhood of a point).

- T_p^*: the dual space to T_p. It is also a linear space of dimension n (the dimension of the manifold). It will be defined as the set of all linear maps from T_p to $\mathcal{R}$. In the presence of a metric (as in Newtonian physics, Special and General Relativity) the latter provides a natural one-to-one correspondence between elements of T_p and its dual. Because of this, we will usually refer to elements of T_p and those in its dual space through this identification by the same name; if there is a need to explicitly distinguish elements in T_p and its dual we will refer to those in the latter as *co-vectors*.

- Upper and lower indices:
  - We denote vectors in the tangent space T_p through lower indices, and *components* of vectors on any given basis by upper indices. For example, for each $\mu = 1, \dots n$, $\partial_u \in T_p$ is a vector, and v^μ denote the components of a vector in T_p [as in Eq.(1)]. For covectors (elements of the dual space T_p^*) the convention is reversed: covectors are labeled by upper indices and their components by lower indices.

- Letter conventions:
  - Greek indices (μ, ν), etc. Denote spacetime indices. Therefore there should be four of them (either running from 0 to 3 or 1 to 4, depending on the source, i.e. book, article or else). If from 0 to 3 (1 to 4), 0 (4) usually refers to the "time" component.
  - Latin indices starting with i. That is, i, j, k, etc denote the spatial components of some spacetime index (i.e. they would typically run from 1 to 3).
  - Latin indices starting with a, as in a, b, c, d refer to the abstract index notation. This is meant to work as if using a coordinate system or basis but emphasizing that the expressions are valid in *any basis*. For example, in the abstract index notation a *vector* itself $v \in T_p$ would be denoted as v^a and a covector $\omega \in T_p^*$ as ω_a. Then the contraction $v^a \omega_a$ is independent of the coordinate system and $v^a \omega_a = v^\mu \omega_\mu$ in any (dual) bases. As another example, raising an lower indices defines tensors of new types which are uniquely defined regardless of the basis used. E.g., we denote by v^b the vector $v \in T_p$ defined as $v^b := g^{ab} v_b$ for any $v_b \in T_p^*$. The resulting vector is independent of the basis used to perform the lowering or raising of indices.

- We will always use bases for the tangent space and its dual which are dual to each other, Eq. (17).

- g: the metric of spacetime and the fundamental entity of General Relativity. At each point of $\mathcal{M}$ it defines the scalar product between any two vectors in T_p. It is a non-positive definite, non-degenerate, scalar product. We refer to it being non-positive definite as saying that the metric of GR (as is that one of Special Relativity) is *Lorentzian*. In contrast, the metric of Newtonian physics is non-negative definite and is therefore said to be *Riemannian*.

- η: the metric of Special Relativity, also called the *Minkowski* metric. In inertial and cartesian coordinates it is simply $\eta = \text{diag}(-1, 1, 1, 1)$

- Unless otherwise stated, we will use *geometric* units, in which Newton's constant and the speed of light are one, $G = c = 1$

3 The geometry of Galilean-Newtonian physics, Special Relativity and General Relativity

3.1 Galilean-Newtonian (GN) physics

In GN physics we have a preferred choice of universal time, and we compute and measure the distance between points in space. For example, if we have two points, labeled by Cartesian coordinates (x^i, y^i, z^i), with $i = 1, 2$, and define the difference between these coordinates as $\Delta x, \Delta y, \Delta z$, we know from Pithagora's theorem that the separation between the them is

$$(\Delta l)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \,. \tag{2}$$

We will sometimes write this kind of expressions in a more succinct form, with a slight abuse of notation,

$$\Delta l^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 \,,$$

remembering that we are not taking the differences of the squares but the squares of the differences.

The simple form of the distance between two points, Eq.(2), is only valid in Cartesian coordinates; for example, if we were to use cylindrical coordinates it would take a very different form. Though it might sound trivial, in Newtonian physics we assume that the distance between two points is independent of the coordinates. For example, if we define a new set in which the z coordinate is rescaled,

$$x' = x, y' = y, z' = z/2$$

then the distance between the two points in these new coordinates is

$$(\Delta l)^2 = (\Delta x')^2 + (\Delta y')^2 + (2\Delta z)^2$$

Furthermore, we can use curvilinear and moving (such as rotating) coordinates and the distance between two points would be the same. Two examples are spherical coordinates (r, θ, ϕ), defined implicitly through

$$x = r\sin\theta\cos\phi \,, y = r\sin\theta\sin\phi \,, z = r\cos\theta \,,$$

and cylindrical ones ρ, θ, z

$$x = \rho\cos\phi \,, y = \rho\sin\phi \,, z = z \,.$$

We can also write the infinitesimal distance between two points (also called the **line element**). In cartesian coordinates this would be

$$(ds^2) = (dx)^2 + (dy)^2 + (dz)^2 \,.$$

Again, for simplicity sometimes we will write the above as $ds^2 = dx^2 + dy^2 + dz^2$. Using the chain rule and imposing that the line element is independent of the coordinates we can write it in any other ones. For example, let's do it for cylindrical coordinates,

$$dx = \frac{\partial x}{\partial \rho} d\rho + \frac{\partial x}{\partial \phi} d\phi = \cos\phi d\rho - \rho\sin\phi d\phi \tag{3}$$

$$dy = \frac{\partial y}{\partial \rho} d\rho + \frac{\partial y}{\partial \phi} d\phi = \sin\phi d\rho + \rho\cos\phi d\phi \tag{4}$$

$$dz = dz \,. \tag{5}$$

Therefore,

$$dl^2 = d\rho^2 + \rho^2 d\phi^2 + dz^2 \,,$$

which we recognize as the standard line element in cylindrical coordinates.

We can introduce the metric g for Newtonian physics, which for now can be thought as a matrix[1] with components, in Cartesian coordinates,

$$g = \mathrm{diag}(1, 1, 1)\,.$$

That is, its components are $g_{ij} = \delta_{ij}$, where δ is the Kronecker delta. It is diagonal in these coordinates and its signature (defined later, for the time being think, again with a grain of salt, of eigenvalues) is $(+1, +1, +1)$. We therefore say that it is a **Riemannian** metric, which is to say that it is positive definite. It is, in particular, invertible, and we denote the components of its inverse by g^{ij} (upper indices instead of lower ones). In Cartesian coordinates the components of the inverse metric equal those of the metric itself, but this is not in general the case. For example, in cylindrical coordinates

$$g^{ij} = \mathrm{diag}(1, \rho^{-2}, 1)\,.$$

If we denote by x^1, x^2, x^3 arbitrary coordinates of space, the line element can now be conveniently written as

$$ds^2 = \sum_{i=1}^{3}\sum_{j=1}^{3} g_{ij}dx^i dx^j = g_{ij}dx^i dx^j\,,$$

where in the second equality we have used the convention of summation over repeated indices.

Newton's equation for the gravitational field $\Phi(t, \vec{x})$ is given by

$$\nabla^2\Phi(t, \vec{x}) = 4\pi G\rho(t, \vec{x})\,, \tag{6}$$

where $\rho(t, \vec{x})$ is the mass density, G Newton's constant and ∇^2 denotes the Laplacian operator. In Cartesian coordinates this is

$$\nabla^2\Phi(t, \vec{x}) = (\partial_x^2 + \partial_y^2 + \partial_z^2)\Phi(t, \vec{x})\,.$$

which we can also write as,

$$\nabla^2\Phi = g^{ij}\nabla_i\nabla_j\Phi = 4\pi G\rho(t, \vec{x})\,,\,. \tag{7}$$

As we will see later, this way of looking at the equation guarantees that it is coordinate independent, or **covariant**. In fact, if one looks for a linear, second order (that is, involving up to second derivatives) covariant partial differential equation for the gravitational field which do not involve extra structure, Eq. (7) is essentially unique.

In (7) ∇ is called a **covariant derivative**, and in cartesian coordinates it is just

$$\nabla_i = \partial_i\,.$$

It is not, however, when using curvilinear coordinates. We will later learn how to compute covariant derivatives of general tensors in arbitrary coordinates and metrics.

Since in GN physics there is no obvious evidence of a preferred direction in space, or a time variation of Newton's constant, we want Eq. (7) to be covariant. Again, using the chain rule we can write down (6) in any coordinate system.

Similarly, the acceleration of a particle in the gravitational field Φ is

$$\vec{a} = -\vec{\nabla}\Phi\,,$$

which in Cartesian coordinates is simply

$$(a_x, a_y, a_z) = -(\partial_x\Phi, \partial_y\Phi, \partial_z\Phi)\,. \tag{8}$$

Notes:

[1] This has to be taken with a grain of salt. As we will see, the metric is a tensor of rank $(0, 2)$, while matrices are tensors of rank $(1, 1)$. In particular, the determinant or eigenvalues of a matrix are basis-independent, but those of a metric *are not*.

- In Eq.(8) we have used lower indices for the components of the acceleration. This is not a coincidence or a random decision, we will come back to it in the next chapter, when we define vector fields and dual vector fields.

- Again, we want the acceleration as a vector to be independent of coordinates, but we will postpone this discussion for a bit. In the next chapter we will learn how the components of vector fields and tensors in general change as we change coordinates.

- The shortest distance between two points in Newtonian geometry is a straight line. The fact that it is a straight line is due to the fact that the space geometry is flat, while the fact that it is shortest (and not largest) is a consequence of the metric being Riemannian. We will discuss these issues in due term.

- The equation (6) is a time-independent equation. In the language of partial differential equations, it is an **elliptic** one[2]. In particular, it implies that if one changes the density of mass at any point, the gravitational potential immediately feels the effect and adjusts itself everywhere to accommodate for this change. This is sometimes called *action at a distance* (this term is also used to refer to the fact that no medium is needed to "transmit" the gravitational effect of a mass on another one) and it implies that in Newton's theory of gravity the speed of propagation is infinite and, in particular, there is no radiation of gravitational energy.

3.2 Special Relativity

In Special Relativity (SR) we now consider a metric of space-time. The metric shares features with the Newtonian case. Namely: i) it is fixed, ii) it is also **flat**. At the same time, it introduces a new key difference. It is not positive definite any more. It is not difficult to convince oneself that in order to build covariant partial differential equations describing the propagation of fields with finite speed of propagation, the metric has to have signature (eigenvalues) $(-1, 1, 1, 1)$. Without motivation we therefore introduce the **Minkowski metric**. From a metric of space to one of space-time we will switch from dl^2 to ds^2, noticing that it is not positive definite and therefore the square is pure historical notation,

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \,, \tag{9}$$

with c the speed of light. Coordinates (t, x, y, z) on which the Minkowski metric takes the above form are called **inertial** and it is one of the assumptions of SR that they exist, of course. Now, much as in the Newtonian case we demanded that dl^2 is independent of the coordinates used, we do the same for the metric (9) (see Problem 4).

Notes:

- The Minkowski metric is usually referred to as η (as opposed to g), and its components as $\eta_{\mu\nu}$. This is mainly so that in the GR case whenever we are referring to the Minkowski metric the symbol representing it should be speak for itself, and save further explanations.

- The signature of the Minkowski metric is $(-1, 1, 1, 1)$, as is the signature (defined in the next chapter more precisely) of any general relativist metric. We call metrics with such signature **Lorentzian**.

[2]The Einstein equations themselves will turn out to have an elliptic sector (the Einstein *constraints*) and an hyperbolic (evolution) one.

In the same way that in the Newtonian case a covariant PDE (partial differential equation) can be written purely in terms of the space metric, without any further structure, the same can be done in SR. For example, the wave equation (also called the massless Klein-Gordon equation) is

$$\Box\Phi = \eta^{\mu\nu}\nabla_\mu\nabla_\nu\Phi \,.$$

Here the symbol $\Box$ is called the D'alambertian (or "box" operator in some neighborhoods), $\eta^{\mu\nu}$ are the components of the inverse metric if seen as a matrix (which equals itself if using inertial coordinates, as was the case in Newtonian geometry if using Cartesian coordinates, but it is in general not true). The symbol ∇ again denotes a covariant derivative, which we have not yet defined but as in the Newtonian case, if using inertial coordinates here it just means a partial derivative, and

$$\Box\Phi(t,x,y,z) = (-c^2\partial_t^2 + \partial_x^2 + \partial_y^2 + \partial_z^2)\Phi(t,x,y,z) \,. \tag{10}$$

This equation is an example of an **hyperbolic** PDE, and it implies a finite speed of propagation (c), see Problem 5.

3.3 General Relativity

At the risk of over-simplifying it, the main differences between General Relativity (GR) and SR are that i) the manifold of at space-time might not be $\mathcal{R}^4$, ii) the metric is not fixed (in either space or time),

$$ds^2 = g_{\mu\nu}(x^\alpha)dx^\mu dx^\nu \,,$$

iii) neither flat, iv) there is in general no natural distinction between space and time (no preferred coordinates or observers, except asymptotically far away from compact sources).

Free-falling observers follow the generalization of straight lines, **geodesics**. The metric is determined by Einstein's equations,

$$G_{ab} = 8\pi G T_{ab} \,, \tag{11}$$

where G_{ab} is a curvature (the Einstein) tensor which is by construction divergence free, $\nabla^a G_{ab} \equiv 0$, which implies "local conservation" of the stress-energy-momentum tensor T_{ab}, $\nabla^a T_{ab} \equiv 0$. Imposing mild conditions on the covariant derivative ∇, it is uniquely determined by the metric of space-time. The Minkowksi metric is a solution of the vacuum ($T_{ab} \equiv 0$) Einstein equations, but it is not the only one: space-time can be curved even in the absence of matter.

3.4 Problems

1. Write the finite distance $(\Delta l)^2$ between two points in Newtonian physics in cylindrical and spherical coordinates.

2. Write the infinitesimal distance $(dl)^2$ between two points in Newtonian physics in spherical coordinates using the chain rule and the components of the inverse metric.

3. Using the chain rule, write the Poisson equation (6) for the Newtonian gravitational potential in spherical and cylindrical coordinates.

4. Using the chain rule, write down the Minkowski metric in terms of "rotating cylindrical coordinates" (t', ρ', ϕ', z'), defined by cylindrical ones (t, ρ, ϕ, z')

$$t' = t\,, = x = \rho\cos(\phi + \omega t)\,, y = \rho\sin(\phi + \omega t)\,, z = z\,,$$

where ω is a constant.

5. Assume plane symmetry in the Klein Gordon equation (66); that is, $\Phi(t, x, y, z) = \Phi(t, x)$. Write the most general solution to such problem in terms of "advanced" and "retarded" solutions. Explain why this type of solutions imply finite speed of propagation. Finally, discuss what kind of initial conditions determine one and only one (existence and uniqueness) solution.

4 Scalars, vectors and tensors

Definition 1. *We denote by $\mathcal{F}$ the set of functions on a manifold $\mathcal{M}$. That is, $f \in \mathcal{F}$ is a map*

$$f : \mathcal{M} \to \mathcal{R}\,.$$

We also call functions *scalars*, by which we mean to emphasize that they are independent of the coordinates used to describe $\mathcal{M}$ (as opposed to the *components* of vectors, for example). Many scalars will come from *contractions* of tensors, defined later (Definition 7), which include the scalar product of vectors.

Example 1. For instance, in Newtonian physics an example of a scalar could the magnitude of the gravitational acceleration for some test particle at some point in space. If we denote the acceleration vector by $a = (a^1, a^2, a^3) = (a^x, a^y, a^z)$ where the coefficients $\{a^i\,, i = 1, 2, 3\}$ are the Cartesian components of the vector, then the square of its magnitude is the scalar product with itself

$$|a|^2 = a \cdot a = \sum_{i=1}^{3}\sum_{j=1}^{3} g_{ij}a^i a^j =: g_{ij}a^i a^j = (a^x)^2 + (a^y)^2 + (a^z)^2\,. \tag{12}$$

where for the time being the g_{ij} coefficients can be thought of as the components of the matrix

$$g = \mathrm{diag}(1, 1, 1)\,. \tag{13}$$

Even though this example is trivial it serves to illustrate several definitions that we will run into in GR. First, the concept of a scalar as a scalar product between vectors. Second the convention of sum over repeated indices (second equality in Eq.(12)). And finally, the concept of a metric; in this case the metric of Galilean-Newtonian physics, Eq.(13).

4.1 The tangent space to a point on a manifold

Since manifolds in general do not need to be linear spaces, the concept of tangential derivatives at a point is in general the most that we can do in terms of equipping them with some differential structure. This leads to the definition of the tangent space to any given point $p \in \mathcal{M}$, defined as follows.

Take an arbitrary set of coordinates $\{x^\mu\,, \mu = 1 \ldots n\}$ (notice the upper indices) for a neighborhood of an arbitrary point $p \in \mathcal{M}$. Then

Definition 2. *The tangent space to p, usually (but not always) denoted by T_p is the linear space of directional derivatives. That is, a vector $v \in T_p$ is a linear map:*

$$v : \mathcal{F} \to \mathcal{R}$$

within the linear space

$$T_p := Span\{\partial_\mu\,, \mu = 1, \ldots, n\}\,.$$

Note:

- The dimension of T_p is clearly n and the set $\{\partial_\mu, \mu = 1, \ldots n\}$ is a basis of it.
- Here and in many places, ∂_μ is used as a shortcut for $\partial/\partial x^\mu$:

$$\partial_\mu := \frac{\partial}{\partial x^\mu}\,.$$

Definition 3. *Given any local coordinate system* $\{x^\mu\,, \mu = 1 \ldots n\}$ *on a neighborhood of a point* $p \in \mathcal{M}$*, the basis (notice the indices down) of* T_p

$$X_\mu := \partial_\mu\,, \mu = 1, \ldots n \tag{14}$$

is called the associated coordinate basis.

Since they form a basis, any vector $v \in T_p$ can be written as

$$v = \sum_{\mu=1}^{n} v^\mu X_\mu = v^\mu X_\mu = v^\mu \partial_\mu\,,$$

where the second equality is just the convention of repeated indices summation (see Section 2).

That is, the action of a vector $v \in T_p$ on a function $f \in \mathcal{F}$ is indeed its derivative at p in the direction of v,

$$v(f) = (v^\mu \partial_\mu f)\,|_p\,.$$

The notation logic might be clear at this point: **we use *upper* indices for the *components* of vectors in T_p and *lower* indices to denote *vectors* themselves** – the convention for covectors is the opposite, as will be discussed in Section 4.2.

The action of vectors in the tangent space (or any other linear space) is defined to be independent of the basis used, which leads to the following.

Transformation rule for vectors in the tangent space:

Since the tangent space is a linear one, the components of any vector in it under a change of basis transform in the usual way. But most of the time we will be using coordinate-based bases, so we can restrict the discussion to how the components of vectors transform under a change of coordinates

$$x^\mu \to x'^\nu\,(x^\mu) \tag{15}$$

and associated change of basis

$$X_\mu \to X'_\mu\,.$$

Recalling the definition of the coordinate basis (14), this only involves using the chain rule,

$$X_\mu = \frac{\partial x'^\nu}{\partial x^\mu} X'_\nu,$$

where we have again used the convention of sum over repeated indices. Unless otherwise stated, we will always use this convention. Therefore the components on the primed basis are

$$v = v^\mu X_\mu = v'^\mu X'_\mu$$

with

$$v'^\nu = v^\mu \frac{\partial x'^\nu}{\partial x^\mu} = v^\mu \partial_\mu (x'^\nu)\,. \tag{16}$$

4.2 The dual tangent space

Given any linear space, there is an associated one, called its dual space. In particular, this is the case for the tangent space T_p and we denote its dual by T^*_p. It is defined as the set of linear maps from T_p to $\mathcal{R}$. That is,

$$T^*_p := \{\text{ all linear maps } \omega : T_p \to \mathcal{R}\}\,.$$

Since by definition T^*_p is a linear space, its members are vectors. But when there is the need to explicitly distinguish between members of T_p and its dual, it is standard notation to refer to the members of the latter as **co-vectors**.

The dimension of T_p^* is also n. One way of seeing this is by noticing that any $\omega \in T_p^*$ is defined by its action on a basis of T_p, which is of dimension n. Another, more explicit, way is to construct an explicit basis for T_p^* , which we need to do anyway.

Definition 4. *Let $\{e_\mu \in T_p, \mu = 1 \dots n\}$ be an arbitrary basis for T_p. Its* **dual basis** $\{e^\mu \in T_p^*, \mu = 1 \dots n\}$ *is defined by imposing their action on the original basis to be*

$$e^\mu(e_\nu) := \delta^\mu{}_\nu \, . \tag{17}$$

Any vector $\omega \in T_p^*$ can therefore be written as

$$\omega = \omega_\mu e^\mu \, , \tag{18}$$

where again we have used the convention of sum over repeated indices.

Observations:

- Notice the opposite convention to that one for elements of T_p: **we label vectors in T_p^* with *upper* indices (as in e^μ) , and their *components* in any basis with *lower* indices** (as the ω_μ coefficients in Eq.(18)). The advantage of these conventions will become apparent soon.

- Notice also the slight (but very standard and very convenient) abuse of notation, in that we are referring to both the elements of a basis of T_p and the members of its dual basis for T_p^* by the same letter (namely, e in the above case). However, from the context it should be clear what we are referring to, in particular because vectors are labeled with a lower index while co-vectors are labeled with an upper index. In fact, later we will extend this standard and very convenient abuse of notation to all vectors and co-vectors.

- Similarly, we use, for example, upper indices to denote both the components of a vector, or a co-vector itself. For example, for a fixed μ, v^μ could be a co-vector or the component of a co-vector. While this might appear ambiguous it is again very convenient when dealing with contractions and as we will see the ambiguity disappears when introducing a metric, since the latter provides a unique mapping between these two interpretations.

- When dealing with the tangent space and its dual, including the case of tensors (defined in Section 4.5), we will always use bases for T_p and T_p^* which are dual to each other.

Any covector ω can be written as in Eq. (18), while any vector $v \in T_p$ can be similarly expressed as

$$v = v^\nu e_\nu \, .$$

Then the action of ω on v is

$$\omega(v) = \omega_\mu e^\mu \left(v^\nu e_\nu\right) = \omega_\mu v^\nu e^\mu(e_\nu) = \omega_\mu v^\mu \, . \tag{19}$$

where the second equality is due to the fact that covectors are linear, and the last equality is because the basis e_ν and e^μ are chosen to be dual to each other [Eq.(17)].

Notes:

- In Eq. (19) ω_μ and v^ν are real numbers, and e^μ, e_ν are covectors and vectors, respectively, despite using the same notation for vectors and components of co-vectors, and viceversa.

- We now see one of the advantages of the upper and lower indices convention: the action of a covector on a vector is simply the sum of the product of their basis components, $\omega(v) = \omega_\mu v^\mu$. This is an example of a *contraction*, which we will define in general later. For the time being notice that the contraction $\omega_\mu v^\mu$ is independent of the bases used (provided they are dual to each other), because it equals $\omega(v)$, and the latter is defined without reference to any basis.

In the Definition 4 $\{e_\mu\}$ is *any* basis of T_p, not necessarily a coordinate-based one. But an important case is the basis dual to a coordinate-based one.

Definition 5. *Given a coordinate-based basis ∂_μ of T_p, its dual basis is denoted by*

$$\{dx^\mu \in T_p^* \,, \mu = 1 \dots n\} \,.$$

As with all dual bases, the action of each member is defined by Eq. (17), which now reads:

$$dx^\mu(\partial_\nu) := \delta^\mu{}_\nu \,.$$

Transformation rules for co-vectors:

Above we described how the components of a vector in a coordinate-based basis change under a transformation of coordinates in the manifold; namely, as in Eq.(16). Similarly, under a coordinate transformation (15) the components of a covector transform according to

$$\omega = \omega_\mu dx^\mu = \omega'_\mu dx'^\mu$$

where

$$\omega'_\nu = \omega_\mu \frac{\partial x^\mu}{\partial x'^\nu} = \omega_\mu \partial'_\nu(x^\mu) \,. \tag{20}$$

4.3 Tensors

A tensor T of type (k, l) is a multilinear (that is, linear on each argument) map that takes k co-vectors and l vectors as arguments and produces a number.

Definition 6. *The outer product $T \otimes T'$ of two tensors T and T' of type (k, l) and (k', l') is a tensor of type $(k + k', l + l')$ defined as follows*

$$\begin{aligned}&(T \otimes T')(w^1, \dots w^k, w^{k+1} \dots w^{k+k'}, v_1, \dots v_l, v_{l+1} \dots v_{l+l'}) := \\ &T(w^1, \dots w^k, v_1, \dots v_l) T'(w^{k+1}, \dots w^{k+k'}, v_{l+1}, \dots v_{l+l'}) \,,\end{aligned}$$

where each of the two terms in the second equation is a real number so their product is just the multiplication of those two numbers.

Definition 7. *A tensor which is the outer product of two tensors is called simple.*

For example, one way to define a tensor is through the outer product of vectors and co-vectors. While not all tensors are of this form, every tensor can be expressed as a sum of outer products of this type. In more detail, let $\{e_\mu\}$ and $\{e^\nu\}$ be dual bases of T_p and T_p^*, respectively, and T an arbitrary tensor of type (k, l). Then T can be written as

$$T = T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} e_1 \otimes \dots \otimes e_k \otimes e^1 \otimes \dots \otimes e^l \,, \tag{21}$$

where as usual the sum over repeated conventions is used. In particular, if using coordinate-based bases, then

$$T = T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} \partial_1 \otimes \dots \otimes \partial_k \otimes dx^1 \otimes \dots \otimes dx^l \,.$$

Notes:

- The convention for the indices (up or down) in the components of T, namely the coefficients $T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}$, follows that one for vectors and covectors.

- Since we always use dual bases (in particular in the above two equations), the coefficients of a tensor are:

$$T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = T(e^1, \ldots, e^k, e_1, \ldots, e_l)\,.$$

  This can be seen by simply evaluating the right hand side of the above equation and using the property (17) of dual bases in Eq. (21).

In Eq.(19) we saw an example of a *contraction* operation for a tensor of a very particular type and noticed that it was independent of the bases (provided they are dual to each other). We now define contraction for an arbitrary tensor which also turns out to be independent of the bases.

Definition 8. *The contraction of a tensor T of type (k, l) is a tensor, let's call it CT, of type $(k-1, l-1)$ defined as follows:*

$$CT = \sum_{\mu=1}^{n} T(\ldots, e^\mu, \ldots, \ldots, e_\mu, \ldots)\,, \tag{22}$$

where $\{e_\mu\}$ and $\{e^\mu\}$ are arbitrary dual bases.

That is, in a contraction two slots are chosen and the above sum performed. The result, of course, depends on what those slots were. From its definition it might appear that the contraction of a tensor also depends on the basis used in the above definition, but it does not and is left as an exercise (problem 5).

In any basis, the components of the contraction are

$$(CT)^{\mu_1\ldots\mu_{k-1}}{}_{\nu_1\ldots\nu_{l-1}} = T^{\mu_1\ldots\sigma\ldots\mu_k}{}_{\nu_1\ldots\sigma\ldots\nu_l}$$

where on the right hand side there is the usual sum over the repeated index (here σ). That is, the act of formally summing over any two indices (one up, one down) of the components of a tensor does not actually depend on in which basis the operation was done. Contractions appear everywhere in GR and are an important tool.

Example 2. Let's revisit the contraction in Eq.(19). Given a vector v and a covector ω, we define a tensor of type $(1, 1)$ by their outer product,

$$T := \omega \otimes v\,.$$

In any basis, if the components of ω and v are ω_μ and v^ν, respectively, the components of T are

$$T_\mu{}^\nu = \omega_\mu v^\nu$$

and its contraction gives a tensor of formally type $(0, 0)$, by which we mean a scalar, defined as

$$CT := T_\mu{}^\mu = \omega_\mu v^\mu$$

where now since there is a repeated index a sum over it is assumed as usual. Since CT equals $\omega(v)$ (see Eq.(19)), it is obvious that it is independent of the basis used to compute it. The same holds for contractions of any tensor.

We often skip the notation of CT and whenever there is a sum over repeated indices we know that it is a contraction and a tensor of lower type. Thus, we would refer to the contraction of T in the above example simply as $T_\mu{}^\mu$.

Transformation rules for tensors:

This is just a generalization of (16) and (20) for vectors and covectors and only uses the chain rule. The result is:

$$T'^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = \frac{\partial x'^{\mu_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x'^{\mu_k}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial x'^{\nu_1}} \cdots \frac{\partial x^{\nu_l}}{\partial x'^{\nu_l}} T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}\,. \tag{23}$$

4.4 The abstract index notation

In the abstract index notation we denote a vector $v \in T_p$ as v^a (or $b, c, d \ldots$ instead of a) and a co-vector $\omega \in T_p^*$ as v_a. It is like the notation for their *components*, except that the abstract index notation is meant to emphasize operations which are independent of any basis representation. For example, v^a denotes the vector v, not its components. But, for example, we know that contractions are basis-independent. Therefore, $v^a \omega_a$ would be the scalar that in any basis representation would equal $v^\mu \omega_\mu$.

As another, example for each $\mu = 1, 2, 3, 4$, ∂_μ is a local vector field, and in the abstract index notation we would denote them as

$$(\partial_\mu)^a \,, \quad \mu = 1, 2, 3, 4 \,.$$

4.5 The metric tensor

In General Relativity, the metric of spacetime is one of the most (if not the most) fundamental tensors.

In general (whether in GR or not) a metric is defined as a non-degenerate symmetric tensor of type $(0, 2)$. That is, it is a map

$$g : T_p \times T_p \to \mathcal{R}$$

which takes two vectors and returns a number, which is their scalar product. For this reason usually the following notation is used [as in Eq.(12)]: for any two vectors u, v,

$$g(u, v) \equiv u \cdot v \,.$$

The symmetric property means that $g(u, v) = g(v, u)$ for all u, v. This is equivalent to the property that in terms of any basis,

$$g = g_{\mu\nu} e^\mu e^\nu \,,$$

the metric components themselves are symmetric, $g_{\mu\nu} = g_{\nu\mu}$.

The non-degeneracy property means that $g(u, v) = 0 \, \forall \, u \in T_p$ if and only if $v = 0$. This implies that g has an inverse,

$$g^{-1} : T_p^* \times T_p^* \to \mathcal{R}$$

in the sense that its contraction with the metric gives, in any dual bases, the Kronecker delta,

$$(g^{-1})^{\alpha\mu} g_{\mu\beta} = \delta^\alpha_\beta \,.$$

It is standard notation to drop the -1 superscript in g^{-1} and from the context infer whether the metric or its inverse is used. For example, the representation of the inverse on any basis $\{e_\mu\}$ of T_p would be written

$$g = g^{\mu\nu} e_\mu e_\nu \,.$$

Thus, if we refer to components $g^{\mu\nu}$ (with indices up) it implies that we are referring to the inverse metric and the opposite with indices down (the metric). Similarly, in the abstract index notation g_{ab} is the metric and g^{ab} its inverse.

If using a coordinate-based basis then one has

$$g = g_{\mu\nu} dx^\mu \otimes dx^\nu$$

and often the $\otimes$ symbol is dropped and simply

$$g = g_{\mu\nu} dx^\mu dx^\nu$$

written. In a sense that will be made precise later, the metric can be thought of as an infinitesimal line element, and it is therefore usually denoted by ds^2 in Special and General Relativity and dl^2 in Newtonian physics. So for the former,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \, .$$

Notice however, that it does not need to be non-negative definite (it is *not* in Special and General Relativity) so the "square" in ds^2 can be thought of notation more than anything else.

Example 3. The metric of Newtonian physics. The space manifold is $\mathcal{R}^3$ and its metric in Cartesian coordinates is

$$dl^2 = dx \otimes dx + dy \otimes dy + dz \otimes dz = dx^2 + dy^2 + dz^2 \, . \tag{24}$$

That is, $g_{ij} = g^{ij} = \delta_{ij}$ (the Kronecker delta). Notice that here we can suspect why the metric can be thought of as an infinitesimal line element. For example, Eq.(24) can be thought as the infinitesimal version of Pithagoras' theorem,

$$(\Delta l)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \, ,$$

with the finite Δ's replaced by infinitesimal d's and making what it would seem a completely wrong and non-sense manipulation of terminology. Again, we will make this precise later.

Example 4. The Minkowski metric. In Special Relativity the spacetime manifold is $\mathcal{R}^4$ and in inertial, Cartesian coordinates the metric is

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \, .$$

Notice that if again we interpret the above as an infinitesimal line element and making what would seem a terribly wrong manipulation of objects we would have

$$ds^2 = -dt^2 + dl^2 \, , \qquad \text{with } dl^2 = dx^2 + dy^2 + dz^2$$

so for two events in spacetime infinitesimally close from each other (one can of course keep pushing it and think of finite ones replacing d's by Δs) one would have (recall that throughout these notes we are using units in which the speed of light is $c = 1$ unless otherwise noted):

Definition 9. *Light-, time-, and space-like related events:*

- $ds^2 = 0$ *if the "velocity" of a "beam of information" joining these two events travels with the speed of light:* $v^2 := dl^2/dt^2 = 1$*. The two events are said to be* **light-like** *related or along the light cone.*
- $ds^2 < 0$ *if* $v < 1$*. This would correspond to actual, material observers and the events are said to be* **timelike** *related.*
- $ds^2 > 0$ *if* $v > 1$*. The events are said to be* **spacelike** *related. No physical information is thought to travel faster than the speed of light so events which are spacelike related are outside the light cone and* **causally disconnected**.

These definitions remain the same for a general curved space-time, as in GR.

The above apparently non-sense manipulation of objects can actually be made precise and we will do so later. For the moment we limit ourselves to a few preliminary definitions.

Definition 10. *A vector v is said to be spacelike, null, or timelike, if $g(v,v) > 0$, $g(v,v) = 0$ or $g(v,v) < 0$, respectively.*

"Raising" and "lowering" indices

The metric and its inverse define a *one to one* mapping between T_p and its dual. We refer to this mapping as *raising* and *lowering* indices, for reasons that should become clear next.

Given any $v \in T_p$, the map

$$L_v := g(\cdot, v) : T_p \to \mathcal{R} \tag{25}$$

is by definition a linear map from T_p to $\mathcal{R}$ and therefore an element of the dual space T_p^*. That is, it takes as an argument a vector $\omega \in T_p$ and returns its scalar product with v:

$$L_v(\omega) = g(\omega, v)\,, \quad \forall\, \omega \in T_p\,.$$

In any given basis $\{e_\mu\,, \mu = 1, \ldots n\}$, we would denote the components of $v \in T_p$ through v^μ,

$$v = v^\mu e_\mu\,,$$

and in the abstract index notation we would denote the vector itself by v^a.

With a slight but very convenient abuse of notation we denote the map $L_v \in T_P^*$ defined for any fixed $v \in T_p$ by Eq. (25) also as v (!!!!). That is, we write,

$$v(\omega) := L_v(\omega) = g(\omega, v) = g_{ab} v^a \omega^b \quad \forall \omega \in T_p\,.$$

This abuse of notation might seem (and probably is at the beginning) very confusing but once you get used to it it is extremely convenient. There is no ambiguity, since the mapping (25) is one to one (because the metric is by definition non-degenerate).

If we see v as an element of T_p we would denote it in the abstract index notation as v^a. If we see it – under the above mapping – as an element of T_p^* we would denote it as v_a. In any dual bases we have, when seen as a vector,

$$v = v^\mu e_\mu$$

and as a covector

$$v = v_\mu e^\mu\,,$$

where

$$v_\mu := v^\nu g_{\mu\nu}\,. \tag{26}$$

It is easy to see that all this is consistent. For example,

$$g(\omega, v) = g_{\mu\nu} \omega^\mu v^\nu = \omega_\nu v^\nu$$

where the last equality follows from the definition of the lowercase components ω_μ, Eq. (26) .

Similarly, if we see v as a covector, then for any $\omega = \omega^\mu e_\mu \in T_p$ we have

$$v(\omega) = v_\nu e^\nu \left(\omega^\mu e_\mu\right) = v_\nu \omega^\mu e^\nu e_\mu = v_\nu \omega^\mu \delta^\nu{}_\mu = v_\nu \omega^\nu\,.$$

where the next to last equality comes from the property of the bases being dual to each other.

In the abstract index notation we would write the last equation simply as

$$v(\omega) = \omega_a v^a\,.$$

Similarly, the inverse metric defines, for any fixed $v \in T_p^*$, a unique associated vector in T_p, which we also call v, in the following way. For any $v \in T_p^*$ define

$$L_v^* := g^{-1}(\cdot, v) : T_p^* \to \mathcal{R}\,.$$

That is, L_v^* takes as an argument a covector $\omega \in T_p^*$ and returns its scalar product with v. The following identities should hopefully be clear by now,

$$L_v^*(\omega) = g^{-1}(\omega, v) = \omega^a v_a = \omega^\mu v_\mu = \omega_a v^a = \omega_\mu v^u \quad \forall \omega \in T_p^*,$$

where

$$\omega^\mu := g^{\mu\nu}\omega_\mu \quad , \quad \omega_\mu := g_{\mu\nu}\omega^\mu,$$

and

$$\omega^a := g^{ab}\omega_b \quad , \quad \omega_a := g_{ab}\omega^b.$$

4.6 Problems

1. Find four linearly independent null vectors in Minkowski spacetime.
2. A spacetime has coordinates x^μ ($\mu = 1, 2, 3, 4$) with basis vectors for the tangent space at each point $\partial/\partial x^\mu$ and dx^μ for its dual. What are the values of

$$dx^4\left(\frac{\partial}{\partial x^4}\right), dx^2\left(\frac{\partial}{\partial x^3}\right), \left(\frac{\partial}{\partial x^4}\right)\cdot\left(\frac{\partial}{\partial x^1}\right), dx^4 \cdot dx^1?.$$

   In the last two expressions a dot means the scalar product with respect to spacetime metric (which is arbitrary in this problem).
3. Prove that the 2-dimensional metric space described by

$$ds^2 = dv^2 - v^2 du^2 \tag{27}$$

   is just the flat 2-dimensional Minkowski space usually described by

$$ds^2 = -dt^2 + dx^2. \tag{28}$$

   Do so by finding a coordinate transformation $x(v, u), t(v, u)$ (or by showing that such a transformation exists) which brings the metric given by Eq. (27) into the form (28).
4. Using the transformation properties of metrics as tensor fields, show that the metric of Newtonian physics, which in Cartesian coordinates is

$$dl^2 = dx^2 + dy^2 + dz^2$$

   when expressed in spherical coordinates, defined through

$$x = r\sin\theta\cos\phi\,, y = r\sin\theta\sin\phi\,, z = r\cos\theta,$$

   takes the form

$$dl^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2. \tag{29}$$

   Compare the result with that one of Problem 2 of Section 3.
5. Prove that the operation of contraction is independent of the choice of basis. That is, the right hand side of Eq. (22) is independent of the choice of $\{e_\mu\}$.
6. Prove that the dual of the dual of the tangent space at any given point p is itself. That is,

$$\left(T_p^*\right)^* = T_p.$$

7. Show that under the natural mapping between T_p and its dual in the presence of a metric, any vector $v^a \in T_p$ or co-vector $v_a \in T_p$,

$$((v^a)^*)^* = v^a \quad , \quad ((v_a)^*)^* = v_a \,.$$

8. Let V be an n-dimensional vector space and g a metric on it. Show (for example, through a Gram-Schmidt orthonormalization procedure) that one can always choose an orthonormal basis $v_1, \ldots v_n$ of V, i.e. a basis such that $g(v_\alpha, v_\beta) = \pm\delta\alpha\beta$.

   Show that the signature, i.e. the numbers of plus and minus, is independent of the choice of the orthonormal basis.

9. The metric of special relativity is

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \,.$$

   Find the components, $g_{\mu\nu}$ and $g^{\mu\nu}$ of the metric and its inverse in "rotating cylindrical coordinates", defined by

$$\begin{aligned} t' &= t \\ x' &= (x^2+y^2)^{1/2}\cos(\phi - \omega t) \\ y' &= (x^2+y^2)^{1/2}\sin(\phi - \omega t) \\ z' &= z \,. \end{aligned}$$

   where $\tan\phi = y/x$.

10. (Problem 6 of Carroll rephrased) Consider a three dimensional manifold with coordinates x^i, $i = 1, 2, 3$, with $x^1 = x$, $x^2 = y$, $x^3 = z$, and let p be the point with coordinates $(x, y, z) = (1, 0, -1)$. Consider the following curves passing through p (λ, μ and σ denote parameterizations of the different curves),

$$\begin{aligned} x^i(\lambda) &= \big(\lambda, (\lambda-1)^2, -\lambda\big) \\ x^i(\mu) &= (cos\mu, \sin\mu, \mu - 1) \\ x^i(\sigma) &= \big(\sigma^2, \sigma^3 + \sigma^2, \sigma\big) \,. \end{aligned}$$

   - Calculate the components of the tangent vectors to these coordinates in the coordinates bases $(\partial_x, \partial_y, \partial_z)$. Do any of these depend on the metric of the considered manifold?.
   - Let $f = x^2 + y^2 - yz$. Calculate $df/d\lambda$, $df/d\mu$ and $df/d\sigma$.

11. Problem 7 of Carroll, Chapter 1.

5 Covariant derivatives, parallel transport and geodesics

5.1 Covariant derivatives

Definition 11. *A covariant derivative ∇ is an operator satisfying the properties (1-5) below such that for any tensor field T of type (k,l) returns one, ∇T, of type $(k,l+1)$. In the abstract index notation we write*

$$\nabla T = \nabla_c T^{a_1 \dots a_k}{}_{b_1 \dots b_l} \,.$$

1. *Linearity:*
$$\nabla(aT + bS) = a\nabla T + b\nabla S$$
for al $a, b \in \mathcal{R}$, and T, S tensors of (the same, so that the sum is defined) type.

2. *Leibnitz rule: if A, B are tensors (not necessarily of the same type), then*
$$\nabla(AB) = (\nabla A)B + A(\nabla B) \,.$$

3. *Commutativity with contractions:*
$$\nabla_d (A^{a_1 \dots c \dots a_k}{}_{b_1 \dots c \dots b_l}) = \nabla_d A^{a_1 \dots c \dots a_k}{}_{b_1 \dots c \dots b_l} \,.$$

4. *Consistency with tangent vectors as directional derivatives (this is equivalent as saying that the covariant derivative of a scalars is just the partial derivative with respect to any local coordinate system, $\nabla_\mu f = \partial_\mu f$):*
$$\forall f \in \mathcal{F} \text{ and } t \in T_p \,, \quad t(f) = t^a \nabla_a f \,.$$

5. *Torsion free:*
$$\forall f \in \mathcal{F} \quad \nabla_a \nabla_b = \nabla_b \nabla_a f \,.$$

   Note: *there are theories of gravity which are not torsion-free, but this involves adding additional structure (fixing the torsion) to the space-time, beyond that one of just a metric.*

In essence ∇ is like a standard derivative but such that ∇T transforms as a tensor. It is easy to see if we defined ∇ to be the partial derivatives of the tensor components with respect to an arbitrary local coordinate system, the result would satisfy all of the above properties but would not transform as a tensor except when T is a scalar. However, consider an arbitrary but fixed local coordinate system $\{x^\mu\}$ and the associated basis for the tangent space and its dual. Define a covariant derivative such that in that basis its components are the standard partial derivatives,

$$\nabla_\alpha T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = \partial_\alpha T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} \,,$$

and, in any other one, what they have to be so that ∇T transforms as a tensor; namely, following the rule (23). One can easily see that this definition does satisfy properties (1-5). We schematically denote such covariant derivative by ∂_a, remembering that it depends on the coordinate system used to define it and only in that system does it correspond to partial derivatives of the tensor components.

Without proof we shall state the following:

- If ∇ and $\tilde{\nabla}$ are any two covariant derivatives, then their difference, when acting on any dual tensor field ω is of the form
$$\nabla_a \omega_b = \tilde{\nabla} \omega_b - C^c{}_{ab} \omega_c \,,$$

where C is a tensor called the Christoffel symbol. That is, the difference between two covariant derivatives is not a differential operator but an *algebraic* one.

Similarly (notice the change of sign),

$$\nabla_a t^b = \tilde{\nabla} t^b + C^b{}_{ac} t^c ,$$

and, in general,

$$\nabla_a T^{b_1 \dots b_k}{}_{c_1 \dots c_l} = \tilde{\nabla}_a T^{b_1 \dots b_k}{}_{c_1 \dots c_l} + \sum_i C^{b_i}{}_{ad} T^{b_1 \dots d \dots b_k}{}_{c_1 \dots c_l} - \sum_j C^d{}_{ac_j} T^{b_1 \dots b_k}{}_{c_1 \dots d \dots c_l} ,$$

where i and j denote the locations of the index d on each term above.

In other words, C defines a covariant derivative. Conversely, if $\tilde{\nabla}$ is a covariant derivative then ∇ as defined above is another covariant derivative for any tensor field C of type $(1, 2)$ such that $C^c{}_{ab} = C^c{}_{ba}$.

Note:

- If $\tilde{\nabla}_a = \partial_a$ then C is denoted by Γ. For example,

$$\nabla_a t^b = \partial_a t^b + \Gamma^b{}_{ac} t^c ,$$

and, in the coordinate system used to define ∇_a,

$$\nabla_\mu t^\nu = \partial_\mu t^\nu + \Gamma^\nu{}_{\mu\sigma} t^\sigma .$$

5.2 Curves and parallel transport

On manifolds the tangent spaces at different points are not related in any natural manner. One way of identifying them is through the concept of parallel transporting vectors along curves joining any two points. Such identification in general depends on the curve and there is not much that one can do about it. Still (or rather, because of it), the concept of parallel transport is very useful as it leads to one way of defining the intrinsic curvature of a manifold.

We start by explicitly discussing the tangent vector field to a curve. Let γ be a curve on a manifold $\mathcal{M}$, i.e. a map

$$\gamma : [a, b] \to \mathcal{M} ,$$

where $[a, b]$ is some interval in $\mathcal{R}$. Thus for each $\lambda \in [a, b]$ (λ is called the *parametrization* of the curve) there is a tangent space $T_{\gamma(\lambda)}$ at $\gamma(\lambda)$ and we define the tangent vector field $t \in T_{\gamma(\lambda)}$ as follows.

Definition 12. *Recall that a vector maps functions into real numbers. We define the action of the tangent vector field t on a function f to be*

$$t(f) := \frac{df}{d\lambda} .$$

If we use coordinates x^μ in a neighborhood of γ and the associated basis for each tangent space $\{\partial_\mu\}$ then we have

$$t(f) = \frac{df}{d\lambda} = \frac{dx^\mu}{d\lambda} \frac{\partial f}{\partial x^\mu} = \frac{dx^\mu}{d\lambda} \partial_\mu (f) .$$

Therefore the components of the tangent vector on any coordinate-based basis are

$$t^\mu = \frac{dx^\mu}{d\lambda} .$$

Notice that the norm of t^a (though not its direction) does depend on the parametrization of the curve.

We will assume that at no point all the components $dx^\mu/d\lambda$ vanish, i.e. that the tangent is not identically zero.

Definition 13. *A vector field v^a is parallel transported along a curve γ with tangent vector t^a, with respect to a covariant derivative ∇, if*

$$t^a \nabla_a v^b = 0 \,. \tag{30}$$

That is, if its derivative of the vector field in the direction of the curve vanishes.

Definition 14. *A curve is spacelike, null or timelike if its tangent vector at every point is spacelike, null or timelike.*

Definition 15. *The length of a spacelike curve γ between two points $\gamma(\lambda_i), \gamma(\lambda_f)$ is*

$$l := \int_{\gamma(\lambda_i)}^{\gamma(\lambda_f)} \left(g_{ab} t^a t^b\right)^{1/2} d\lambda \,. \tag{31}$$

Similarly, for a timelike curve we define the proper time between two events as

$$\tau := \int_{\gamma(\lambda_i)}^{\gamma(\lambda_f)} \left(-g_{ab} t^a t^b\right)^{1/2} d\lambda \,. \tag{32}$$

The length of a null curve is zero.

Notice that these definitions are independent of the parametrization. That is if we have $\lambda' = \lambda'(\lambda)$ with $d\lambda'/d\lambda \neq 0$ then, in any coordinate system x^μ

$$\frac{dx^\mu}{d\lambda'} = \frac{dx^\mu}{d\lambda}\frac{d\lambda}{d\lambda'}$$

and

$$\int_{\gamma(\lambda_i)}^{\gamma(\lambda_f)} \left(g_{\mu\nu}\frac{dx^\mu}{d\lambda'}\frac{dx^\nu}{d\lambda'}\right)^{1/2} d\lambda' = \int_{\gamma(\lambda'_i)}^{\gamma(\lambda'_f)} \left(g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{1/2} d\lambda \,,$$

provided the endpoints are the same, of course, $\gamma(\lambda'_i) = \gamma(\lambda_i), \gamma(\lambda'_f) = \gamma(\lambda_f)$.

5.3 Metric-compatible covariant derivative

A natural way to choose a (as we will see, unique) natural covariant derivative ∇ is by requiring that the "angle" between any two vectors which are parallel transported along any curve remains constant. As we will now see, this implies that the covariant derivative of the metric is zero. So, we require that for any curve γ with tangent t and any vector fields $v, \omega \in T_\gamma$ which are parallel transported along γ,

$$t^c \nabla_c (v^a \omega^b g_{ab}) = 0 \,.$$

Using the Leibnitz rule and the definition of parallel transport, Eq.(30), we then get

$$t^c v^a \omega^b \nabla_c g_{ab} = 0 \,, \quad \forall \,, t, v, \omega \,.$$

That is,

$$\nabla_c g_{ab} = 0 \,. \tag{33}$$

Definition 16. *Given a metric g on a manifold, a covariant derivative ∇ is said to be metric-compatible if Eq. (33) holds.*

We shall state without proof the following:

Theorem 1. *Given a metric g on a manifold there is a unique metric-compatible covariant derivative ∇. In terms of any other covariant derivative $\tilde{\nabla}$, the respective Christoffel symbols are given by*

$$C^c{}_{ab} = \frac{1}{2} g^{cd} \left(\tilde{\nabla}_a g_{bd} + \tilde{\nabla}_b g_{ad} - \tilde{\nabla}_d g_{ab} \right) . \tag{34}$$

From now on we will always use metric-compatible derivatives.

Notes:

- Notice that, as they should, the Christoffel symbols (34) are symmetric in the last two indices, $C^c{}_{ab} = C^c{}_{ba}$.
- In practice we will always be using $\tilde{\nabla}_a = \partial_a$, i.e. standard partial derivatives. In that case we usually write $\Gamma^c{}_{ab}$ instead of $C^c{}_{ab}$. That is,

$$\Gamma^c{}_{ab} := \frac{1}{2} g^{cd} \left(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab} \right) .$$

  The symbols $\Gamma^c{}_{ab}$, beyond the *Christoffel symbols*, go under a variety of names, including *connection coefficients.*

5.4 Geodesics

In Newtonian physics there is a clear, intuitive notion of straight lines, and they turn out to minimize the distance between two points in space. The concept of geodesic is a generalization of a straight curve on a curved manifold and as we will, see, they are also local extrema of distance or proper separation between two points. The concept of a "straight" curve not surprisingly leads to defining a geodesic as a curve for which its tangent vector is parallel transported along itself.

Definition 17. *Let*

$$\gamma : [a, b] \to \mathcal{M}$$

be a curve on a manifold $\mathcal{M}$ with non vanishing tangent vector t^a. We say that γ is a geodesic with affine parametrization if t^a is parallel transported,

$$t^a \nabla_a t^b = 0 . \tag{35}$$

Notes

- Notice that geodesics with affine parametrization remain timelike, null or spacelike. That is, cannot change from one type to another, because (35) implies that the norm of their tangent remains constant along them,

$$\frac{d}{d\lambda}(t^a t_a) = t^c \nabla_c (t^a t^b g_{ab}) = t^c t^a t^b \nabla_c g_{ab} + g_{ab} t^c \nabla_c (t^a t^b) = g_{ab} t^c \nabla_c (t^a t^b) = 2 g_{ab} t^a t^c \nabla_c t^b = 0 ;$$

  where the first equality is just Leibnitz rule, the second one uses the fact that our covariant derivative is metric compatible, the third one Leibnitz rule again and the last one the definition of geodesic as in Eq. (35).

- Therefore, without loss of generality we can (and usually will) choose the parametrization such that $t^a t_a = 1, 0, -1$, depending on whether the geodesic is spacelike, null or timelike, respectively.

Existence of geodesics with affine parametrizations

Expanding the geodesic equation (35), and using a coordinate system,

$$\begin{aligned} 0 &= t^\mu \left(\partial_\mu t^\nu + \Gamma^\nu{}_{\mu\sigma} t^\sigma\right) = \frac{dt^\nu}{d\lambda} + \Gamma^\nu{}_{\mu\sigma} t^\sigma t^\nu \\ 0 &= \frac{dx^\nu}{d\lambda^2} + \Gamma^\nu{}_{\mu\sigma} \frac{dx^\sigma}{d\lambda}\frac{dx^\nu}{d\lambda} \,. \end{aligned} \tag{36}$$

A simple example is that one of Newtonian physics. If using Cartesian coordinates all the Christoffel symbols vanish and the previous equation has as solutions (switching to latin indices to explicitly denote that we are working in *space*)

$$x^i(\lambda) = \lambda \frac{dx^i}{d\lambda}(\lambda = \lambda_0) + x^i(\lambda = \lambda_0) \quad i = 1, 2, 3\,; \tag{37}$$

that is, the straight lines we are used to.

What can be said in general?. Equation (36) is a system of n second order ordinary differential equations for the coordinates $x^\nu(\lambda)$ of the geodesic. As such, for any initial point $x^\nu(\lambda = \lambda_0)$ and tangent vector components $(dx^\nu/d\lambda)(\lambda = \lambda_0)$ there is one and only one solution to Eq. (36). That is, through every point and for any given tangent vector components there is a unique affinely parametrized geodesic passing through it with such tangent.

(Non) Uniqueness of affine parametrizations

The previous note bears the question of how much freedom there is in affine parametrizations. That is, in rescaling the tangent vector field t^a to a geodesic while still satisfying Eq (35). As we will now see, they are all related to each other through linear transformations. That is, if λ and λ' are both parametrizations of a curve γ,

$$t^\mu = \frac{dx^\mu}{d\lambda}\,,\; t'^\mu = \frac{dx^\mu}{d\lambda'} \quad \text{such that} \quad t^a \nabla_a t^b = 0 = t'^a \nabla_a t'^b$$

then

$$\lambda = a\lambda' + \lambda_0$$

for some constants a, λ_0.

The proof is as follows:

$$t^a \nabla_a t^b = 0 = t'^a \nabla_a t'^b = \alpha t^a \nabla_a (\alpha t^b)\,,$$

where

$$\alpha = d\lambda/d\lambda'\,. \tag{38}$$

Using the Leibnitz rule and again the property that λ is assumed to be an affine parameter,

$$0 = \alpha t^a \nabla_a(\alpha t^b) = \alpha t^a \left(\alpha \nabla_a t^b + t^b \nabla_a \alpha\right) = \beta t^b\,, \tag{39}$$

with

$$\beta := \alpha t^a \nabla_a \alpha\,. \tag{40}$$

Since $t^b \neq 0$, Eq.(39) implies $\beta = 0$. Since the transformation between the two parametrizations is implicitly assumed to be invertible, $\alpha \neq 0$ and Eq.(40) implies that α is constant along the geodesic,

$$0 = t^a \nabla_a \alpha = \frac{d}{d\lambda}\alpha\,, \quad \alpha = a \;\text{ for some constant } a.$$

Recalling that α is the Jacobian between the two parametrizations, Eq. (38),

$$\lambda = a\lambda' + \lambda_0 \,, \text{ for some constant } \lambda_0 \,.$$

Non-affinely parametrized geodesics

We defined affinely parametrized geodesics as those curves where the tangent not only remains "pointing in the same direction" but also remains constant in norm. The more general definition of a geodesic relaxes the latter:

Definition 18. *Let*

$$\gamma : [a,b] \to \mathcal{M}$$

be a curve on a manifold $\mathcal{M}$ with non vanishing tangent vector t^a. We say that γ is a geodesic if

$$t^a \nabla_a t^b = \alpha t^b \qquad \textit{for some function } \alpha. \tag{41}$$

It can be seen, and is left as a problem, that for any geodesic satisfying (41) there is a reparametrization $\lambda' = \lambda'(\lambda)$ of $\gamma(\lambda)$ such that (35) holds. The proof essentially involves explicitly writing down the coordinate transformation as an ordinary differential equation and noticing that it has a (non-unique, since we now know that affine parametrizations are not unique) solution.

Geodesics as extrema of length/proper time

We already pointed out that for the standard line element

$$ds^2 = dx^2 + dy^2 + dz^2 \,,$$

geodesics are straight lines, and we know that they minimize the distance between two points.

As we will now see, in the general case the length or proper time of geodesics is a local extrema (though not necessarily a minimum or maximum). We show it for the case of spacelike geodesics, the timelike case follows the same identical steps.

To be precise, we will consider a curve

$$\gamma_0 : [a,b] \to \mathcal{R}$$

and smooth deformations γ_z (parametrized by z), with the endpoints fixed,

$$\gamma_0(a) = \gamma_z(a)\,, \gamma_0(b) = \gamma_z(b) \quad \forall z \,, \text{with } z \text{ sufficiently small}\,,$$

their lengths

$$l(z) := l(\gamma_z)\,,$$

and we will show that the condition

$$\left.\frac{dl(z)}{dz}\right|_{z=0} = 0$$

is equivalent to the geodesic equation for γ_0.

In order to keep the notation from becoming cumbersome, though, we will not write explicitly the dependence on z but instead write

$$l = \int_{\gamma(a)}^{\gamma(b)} (g_{ab}t^a t^b)^{1/2} d\lambda$$

and compute an infinitesimal variation δl when μ is varied. In addition, without loss of generality we will assume $[a,b] = [0,1]$. Then we have

$$\delta l = \int_{\gamma(0)}^{\gamma(1)} (g_{ab}t^a t^b)^{-1/2} \frac{1}{2} \left[(\delta g_{ab}) t^a t^b + 2 g_{ab} t^a \delta t^b \right] d\lambda \,.$$

Next, we set a coordinate system x^μ and write

$$\delta l = \int_{\gamma(0)}^{\gamma(1)} (g_{\mu\nu} t^\mu t^\nu)^{-1/2} \frac{1}{2} \left[(\delta g_{\mu\nu}) t^\mu t^\nu + 2 g_{\mu\nu} t^\mu \delta t^\nu \right) d\lambda \tag{42}$$

and

$$t^\mu = \frac{dx^\mu}{d\lambda}$$

$$\delta g_{\mu\nu} = \frac{\partial g_{\mu\nu}}{\partial x^\sigma} \delta x^\sigma \,,$$

$$\delta t^\nu = \delta \left(\frac{dx^\nu}{d\lambda} \right) = \frac{d(\delta x^\nu)}{d\lambda} \,.$$

In addition, without loss of generality we assume that the parametrization is such that $g_{\mu\nu} t^\mu t^\nu = 1$. The final step is to integrate by parts the second term in Eq. (42),

$$\int_{\gamma(0)}^{\gamma(1)} g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{d(\delta x^\nu)}{d\lambda} d\lambda = -\int_{\gamma(0)}^{\gamma(1)} \frac{d}{d\lambda} \left(g_{\mu\nu} \frac{dx^\mu}{d\lambda} \right) \delta x^\nu d\lambda + g_{\mu\nu} \frac{dx^\mu}{d\lambda} \delta x^\nu \Big|_{\gamma(0)}^{\gamma(1)} = -\int_{\gamma(0)}^{\gamma(1)} \frac{d}{d\lambda} \left(g_{\mu\nu} \frac{dx^\mu}{d\lambda} \right) \delta x^\nu d\lambda$$

where the boundary terms cancel because we are holding the endpoints of the curve fixed.

Putting all the pieces together and renaming some muddy indices,

$$\delta l = \int_{\gamma(0)}^{\gamma(1)} \left[-\frac{d}{d\lambda} \left(g_{\mu\nu} \frac{dx^\mu}{d\lambda} \right) + \frac{1}{2} \frac{\partial g_{\mu\sigma}}{\partial x^\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\sigma}{d\lambda} \right] \delta x^\nu d\lambda.$$

and

$$\delta l = 0 \quad \forall \quad \delta x^{nu}$$

if and only if

$$0 = -\frac{d}{d\lambda} \left(g_{\mu\nu} \frac{dx^\mu}{d\lambda} \right) + \frac{1}{2} \frac{\partial g_{\mu\sigma}}{\partial x^\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\sigma}{d\lambda} \,. \tag{43}$$

Using the Leibnitz rule,

$$\frac{dg_{\mu\nu}}{d\lambda} = \frac{\partial g_{\mu\nu}}{\partial x^\beta} \frac{dx^\beta}{d\lambda}$$

and the definition (5.3) of the Christoffel symbols, this is exactly the geodesic equation with affine parametrization 35.

As a consequence of geodesics being extrema of the distance or proper length, they can be obtained from the Lagrangian

$$\mathcal{L} = g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}$$

through the familiar from classical mechanics Euler-Lagrange equations:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \right) = \frac{\partial \mathcal{L}}{\partial x^\mu} \,, \tag{44}$$

where $\dot{x}^\mu = dx^\mu / d\lambda$.

5.5 Problems

1. Consider the metric of a 2-sphere with unit radius,

$$ds^2 = d\theta^2 + \sin^2\theta d\phi^2 \,. \tag{45}$$

   Compute the connection coefficients for the covariant derivative associated with this metric.

2. For the 2-dimensional metric $ds^2 = (dx^2 - dt^2)/t^2$ find all the connection coefficients and find all timelike geodesic curves.

3. Consider the Minkowski metric in $1+1$ dimensions,

$$ds^2 = -dt^2 + dx^2$$

   and the curve given by

$$x(\theta) = \cos\theta \,, \; t(\theta) = \sin\theta \,.$$

   That is, the unit circle in the (t, x) plane.

   (a) Write down the tangent vector to it in the $\{\partial_t, \partial_x\}$ basis.

   (b) Determine which segments of the curve are timelike, spacelike, and null.

   (c) For each segment which is timelike, spacelike and null, write down an expression for the length of that segment.

4. The Laplacian operator in Newtonian physics is defined as

$$\nabla^2 := g^{ab}\nabla_a\nabla_b$$

   where g is the Newtonian metric, which in Cartesian coordinates is $g = \text{diag}(1,1,1)$, and ∇_a is its associated covariant derivative.

   Show that in Cartesian coordinates (this is quite trivial, so if you are looking for hidden questions or subtleties, there are none)

$$\nabla^2 = \partial_x^2 + \partial_y^2 + \partial_z^2 \,.$$

   Next, using the form of the metric in spherical coordinates, Eq.(29) and by computing its associated connection coefficients, write ∇^2 in these coordinates.

5. The wave or d'Alembertian operator is the equivalent of ∇^2,

$$\Box := g^{ab}\nabla_a\nabla_b$$

   but with g the (Minkowski) metric of special relativity. More precisely, the latter in Cartesian coordinates is $\text{diag}(-1,1,1,1)$. Show that using those coordinates (again, this is quite trivial, so if you are looking for hidden questions or subtleties, there are none)

$$\Box = -\partial_t^2 + \partial_x^2 + \partial_y^2 + \partial_z^2 \,.$$

   Next, by computing the connection coefficients of the Minkowski metric in spherical coordinates, write down the expression of the wave operator in those coordinates.

6. Show that for any geodesic satisfying (41) there is an affine re-parametrization $\lambda' = \lambda'(\lambda)$ of $\gamma(\lambda)$. That is, such that Eq. (35) holds.

6 Curvature, the Einstein equations

In this section we will first introduce the Riemann curvature of any given manifold equipped with a metric and its unique metric-compatible, torsion-free covariant derivative. There are a number of ways of doing so, here we will proceed through the following steps:

1. Define the Riemann tensor through the lack of two covariant derivatives on tensor fields to commute. Interestingly enough, the commutator will turn out to depend *algebraically* (as opposed to differentially) on the tensor fields. That is, will only depend on the pointwise value of the field.

2. Mention that the Riemann tensor characterizes the difference between the initial and final vector at any given point when parallel transported around an infinitesimal closed loop.

3. Present explicit expressions for the Riemann tensor in terms of the metric, Christoffel symbols and derivatives.

4. Discuss the geodesic deviation equation. This relates the “acceleration" between nearby geodesics to the Riemann tensor.

After those steps we will introduce Einstein's equations, which are very simply related to (a contraction of) the Riemann tensor.

6.1 The Riemann tensor as lack of two covariant derivatives to commute

Recall that one of the properties that we demanded of a covariant derivative was that it was torsion free. That is, for any function f, the anti-symmetric part of its second derivative should identically vanish,

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)\, f \equiv 0\,. \tag{46}$$

This is not true for general tensor fields, though, but the antisymmetric part of the second derivative has very interesting properties in describing the curvature of a manifold. Let ω_a be an arbitrary dual vector field, we will show the following:

Theorem 2. *There is a tensor field* ${R_{abc}}^d$ *of type* $(1,3)$*, called the Riemann tensor, such that*

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)\,\omega_a = {R_{abc}}^d\omega_d\,. \tag{47}$$

Notice that Eq. (58) in particular implies that even though $\nabla_a\nabla_b$ is a (second order) differential operator, its anti-symmetric part is only algebraic, as the right hand side of (58) only depends on v_a at the given point. Below we will derive expressions to (58) for general tensor fields, but first we shall prove the above theorem.

Proof: The proof will proceed in two steps. First we prove the following

Lemma 1. *Let* ω_c *be an arbitrary dual vector field and* f *an arbitrary function. Then at any point* p*,*

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)\,(f\omega_c)\big|_p = f(p)\,(\nabla_a\nabla_b - \nabla_b\nabla_a)\,\omega_c\big|_p\,. \tag{48}$$

In order to keep the notation simple we will usually write expressions as (48) just as

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)\,(f\omega_c) = f\,(\nabla_a\nabla_b - \nabla_b\nabla_a)\,\omega_c\,.$$

The proof of (48) only requires the Leibnitz rule and the torsion-free property. Using the Leibnitz rule twice,

$$\begin{aligned}\nabla_a\nabla_b(f\omega_c) &= \nabla_a\left(\omega_c\nabla_b f + f\nabla_b\omega_c\right)\\ &= (\nabla_a\nabla_b f)\omega_c + (\nabla_b f)\nabla_a\omega_c + (\nabla_b\nabla_a f)\omega_c + (\nabla_a f)\nabla_b\omega_c\,.\end{aligned} \tag{49}$$

Swapping the (a, b) indices,

$$\nabla_b \nabla_a (f\omega_c) = (\nabla_b \nabla_a f)\omega_c + (\nabla_a f)\nabla_b \omega_c + (\nabla_a \nabla_b f)\omega_c + (\nabla_b f)\nabla_a \omega_c \,. \tag{50}$$

Then (48) follows from substracting Eq. (50) from Eq. (52) and using the torsion-free property (46).

Even though we will explicitly prove Theorem 2, Lemma 1 already strongly hints that it is true, for it is essentially saying that we can change the behavior of any dual vector field in a neighborhood of any point (by multiplying it by a function) and the antisymmetric double derivative will only depend algebraically on the way we changed the vector field, and only through its value at p. But let's continue with the proof. We need another

Lemma 2. *If ω_c and ω'_c are any two vector fields which agree at some point p, then at that point*

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)\, \omega_c = (\nabla_a \nabla_b - \nabla_b \nabla_a)\, \omega'_c \,.$$

That is, $(\nabla_a \nabla_b - \nabla_b \nabla_a)\, \omega_c$ only depends on the value of ω_c at p, as we have been anticipating. The proof of this lemma is as follows. Take any dual basis vector field $\{e_b^{(\alpha)}\,,\ \alpha = 1, \ldots, n\}$ defined in a neighborhood of p. That is, (α) is not a tensor index (that's why it is between parenthesis) but each α labels a dual vector field. And by assumption the set of the n fields $\{e_b^{(1)}, \ldots, e_b^{(n)}\}$ is, at each point in a neighborhood of p, a basis for the dual tangent space. Then the difference $\omega_b - \omega'_b$ can be expanded in terms of them,

$$\omega_b - \omega'_b = \sum_{\alpha=1}^{n} f_{(\alpha)} e_b^{(\alpha)}$$

where all the functions $f_{(\alpha)}$ vanish at p because by assumption ω_b and ω'_b vanish there,

$$f_{(\alpha)}(p) = 0 \quad \text{for} \;\; \alpha = 1, \ldots, n. \tag{51}$$

Then

$$\begin{aligned} \nabla_a \nabla_b (\omega_c - \omega'_c) &= \nabla_a \nabla_b \left(\sum_{\alpha=1}^{n} f_{(\alpha)} e_b^{(\alpha)} \right) = \sum_{\alpha=1}^{n} \nabla_a \left(f_{(\alpha)} \nabla_b e_c^{(\alpha)} + e_c^{(\alpha)} \nabla_b f_{(\alpha)} \right) \\ &= \sum_{\alpha=1}^{n} \left(f_{(\alpha)} \nabla_a \nabla_b e_c^{(\alpha)} + (\nabla_a e_c^{(\alpha)}) \nabla_b f_{(\alpha)} + (\nabla_a f_{(\alpha)}) \nabla_b e_c^{(\alpha)} + e_c^{(\alpha)} \nabla_a \nabla_b f_{(\alpha)} \right) , \end{aligned}$$

where we have just used linearity of the covariant derivative and the Leibnitz rule. Interchanging the (a, b) indices, using the torsion free property (46) and Eq. (51), we arrive to

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)\, (\omega_c - \omega'_c)|_p = 0 \,.$$

Having proved Lemma 2, the proof of Theorem 2 simply involves noticing that at any point p the map:

$$\omega_c \to (\nabla_a \nabla_b - \nabla_b \nabla_a)\omega_c$$

is a linear map from a dual vector to a tensor T of type $(0, 3)$. Therefore by definition there is a tensor R of type $(1, 3)$ such that

$$T = R(\omega) \,,$$

which is another way of expressing Eq. (58).

Notes:

- If the Riemann tensor vanishes in any open set, the manifold there is said to be **flat**.
- Notice that Eq. (58) automatically implies that the Riemann tensor is antisymmetric in its first two indices,

$$R_{abc}{}^{d} = -R_{bac}{}^{d} .$$

- The convention for the order of the indices in the Riemann tensor changes from reference to reference. For example, our convention agrees with that one of Wald's and not that one of Carroll. The relationship with the latter is the following:

$$R_{abc}{}^{d}(\textbf{here}) = R^{c}{}_{dab}(\textbf{Carroll}) .$$

- Sometimes one would use shortcuts such as

$$\nabla_{ab} := \nabla_a \nabla_b . \tag{52}$$

- Given a type $(0,2)$ tensor T (similar definitions hold for arbitrary ones), its symmetric and antisymmetric parts are, respectively,

$$T_{(ab)} := \frac{1}{2}\left(T_{ab} + T_{ba}\right) \tag{53}$$

$$T_{[ab]} := \frac{1}{2}\left(T_{ab} - T_{ba}\right) \tag{54}$$

$$\tag{55}$$

  and one has $T_{(ab)} = T_{(ba)}, T_{[ab]} = -T_{[ba]}$, and $T_{ab} = T_{(ab)} + T_{[ab]}$.
  In view of this one could (and usually does) write expression such as

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) = 2\nabla_{[ab]} .$$

  In other places (such as Carroll's) the following notation is instead used,

$$[\nabla_a, \nabla_b] := (\nabla_a \nabla_b - \nabla_b \nabla_a) \tag{56}$$

  The previous notation is **not to be confused** with the commutator between two vector fields, which we will define below in Sec. **??** and for which it is standard to use a very similar notation. In these notes we will not use notation such as (56) (in particular because $\nabla_a \nabla_b$ is not a tensor product of two type $(0,1)$ tensors ∇_a and ∇_b), but we will use standard notation such as those of Eqs.(52,54,54).

- It can be seen that the following, called the Bianchi identity, holds,

$$\nabla_{[a} R_{bc]d}{}^{e} = 0 , \tag{57}$$

  where the brackets indicate antisymmetrization in a, b, c.

6.2 Properties

In any coordinate system, the Riemann tensor is given by

$$R_{\mu\nu\rho}{}^{\sigma} = \partial_\nu \Gamma^{\sigma}{}_{\mu\rho} - \partial_\mu \Gamma^{\sigma}{}_{\nu\rho} + \Gamma^{\alpha}{}_{\mu\rho} \Gamma^{\sigma}{}_{\alpha\nu} - \Gamma^{\alpha}{}_{\nu\rho} \Gamma^{\sigma}{}_{\alpha\mu} , \tag{58}$$

with the Ricci tensor obtained by summing (contracting) over ν and σ.

Notice from (58) that if a metric is constant in any open set, then its Riemann tensor there vanishes and the space-time is flat. It can be seen (see Problem 1) that the metric is the Minkowski one. Conversely, it is more difficult but it can be seen that if the Riemann tensor vanishes in an open set then the metric is also the Minkowski one.

6.3 Einstein's vacuum equations

By vacuum equations we mean in the absence of matter fields. That is, in General Relativity one can have a non-trivial gravitational field even in the absence of matter. We will introduce matter fields into Einstein's equations a bit below.

Definition 19. *The Ricci tensor R_{ab} is defined by the following contraction of the Riemann tensor $R_{abc}{}^{d}$,*

$$R_{ab} := R_{acb}{}^{c} .$$

It can be seen that the Ricci tensor is symmetric, $R_{ab} = R_{ba}$.

Definition 20. *The Ricci scalar R, in turn, is defined as the trace of the Ricci tensor,*

$$R := R_{a}{}^{a} .$$

Definition 21. *The Einstein tensor G_{ab} is defined as*

$$G_{ab} := R_{ab} - \frac{1}{2} R g_{ab} . \tag{59}$$

Now, finally what you have been waiting for: **the Einstein vacuum equations are**, simply,

$$G_{ab} = 0 . \tag{60}$$

Notes:

- We (and it is standard practice) are using the same letter "R" to denote the Riemann and Ricci tensors and the Ricci scalar. However, there should be no ambiguity from the context, and you can figure out which one is being referred to depending on how many indices each "R" has.
- It is also standard practice to denote, as we are doing in these notes, the Einstein tensor by a capital G (unfortunately, the same letter used for Newton's constant!) and the metric one by a lower case g (i.e. these notation conventions are case sensitive!).
- Taking the trace of Eq.(59), we have (recall that by definition $g_{a}{}^{b} = \delta_{a}{}^{b}$)

  $$G := G_{a}{}^{a} = R - \frac{1}{2} R \delta_{a}{}^{a} = R - 2R = -R$$

  and therefore Eq.(59) can be inverted,

  $$R_{ab} := G_{ab} - \frac{1}{2} G g_{ab} .$$

- In view of the previous observation, if the Einstein equations (60) hold, then

  $$R_{ab} = 0 \tag{61}$$

  and viceversa. The reason for introducing the Einstein tensor will become apparent when we discuss matter fields. For the vacuum case it is enough to consider Eq. (61) as the Einstein equations.

6.4 Matter fields

The stress energy momentum tensor T for any matter field is a symmetric, type $(0, 2)$ tensor, $T_{ab} = T_{ba}$.

- For an observer with 4-velocity u^a, $T_{ab}u^a u^b$ is interpreted as the energy density, i.e. mass density per volume, as measured by that observer. For "normal" matter, $T_{ab}u^a u^b \geq 0$. This is called the **Weak Energy Condition**.
- If x^a is a vector perpendicular to u^a, $v^a u_a = 0$, $-T_{ab}u^a x^b$ is the momentum density of matter in the x^a direction.
- If y^a is another vector perpendicular to both x^a and u^a, $T_{ab}x^a y^b$ is the $x - y$ component of the stress tensor.

6.4.1 Special Relativity

Consider first the case of special relativity, with the Minkowski metric in inertial, Cartesian coordinates,

$$\eta_{ab} = -dt^2 + d\ell^2 \,.$$

Since it has constant coefficients, it satisfies $\partial_a \eta_{bc} = 0$.

Example 5. A perfect fluid in special relativity. The stress-energy momentum tensor is given by

$$T_{ab} = \rho u_a u_b + P(\eta_{ab} + u_a u_b) \,,$$

where ρ is the density and P the pressure.

In special relativity the field equations for matter fields always follow from the conservation law $\partial^a T_{ab} = 0$. For example, for the above example of a perfect fluid, this yields

$$0 = u^a \partial_a \rho + (\rho + P)\partial^a u_a \tag{62}$$

$$0 = (\rho + P)u^a \partial_a u_b + (\eta_{ab} + u_a u_b)\partial^a P \tag{63}$$

In the non-relativistic limit in which $P \ll \rho$, $u^\mu = (1, \vec{v})$, and $|\vec{v}|dP/dt \ll |\vec{\nabla}P|$, these equations become

$$0 = \frac{\partial \rho}{\partial t} + \vec{\nabla}(\rho \vec{v}) \tag{64}$$

$$\rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \vec{\nabla})\vec{v} \right] = -\vec{\nabla}P \,. \tag{65}$$

Equation (64) represents conservation of mass, and Eq. (65) are the Euler equations of fluid dynamics.

As another example of how the matter field equations are given by $\partial^a T_{ab} = 0$ consider now

Example 6. A scalar field Φ with mass m; the stress-energy-momentum tensor is given by

$$T_{ab} = (\partial_a \Phi)(\partial_b \Phi) - \frac{1}{2} g_{ab} \left(\partial^c \Phi \partial_c \Phi + m^2 \Phi^2 \right) \,.$$

It is easy to show that $\partial^a T_{ab} = 0$ in this case gives

$$\nabla^a \nabla_a \Phi = m^2 \Phi \,, \tag{66}$$

which is the expected wave propagation for a massive scalar field (also called the **Klein-Gordon** equation).

Our last example is that one of electromagnetic (EM) fields:

Example 7. The electric and magnetic fields are combined into an EM tensor F_{ab}, which is antisymmetric, $F_{ab} = -F_{ba}$. For an observer with 4-velocity u^a,

$$E_a = F_{ab}u^b$$

is the electric field as measured by u^a. Similarly,

$$B_a = -\frac{1}{2}\epsilon_{ab}{}^{cd}F_{cd}u^b \tag{67}$$

is the magnetic field measured by the same observer. In Eq. (67), ϵ_{abcd} is an example of a *volume element*, which is antisymmetric in all its indices and, having chosen an *orientation*, its only independent component is given by $\epsilon_{0123} = \sqrt{|g|}$, with g the determinant of the metric. It is not a tensor, because it does not transform as such. We will not elaborate on it, at least for the time being. For the moment it suffices to say that, for example, in case of Minkowski space-time in inertial, Cartesian, coordinates, Eq. (67) means $\epsilon_{0123} = 1, -\epsilon_{1023} = \epsilon_{0132}$ and similarly for all the other permutations.

The special relativistic EM stress-energy-momentum tensor is

$$T_{ab} = \frac{1}{4\pi}\left[F_{ac}F_b{}^c - \frac{1}{4}\eta_{ab}F_{de}F^{de}\right] .$$

6.4.2 General relativity

In order to generalize matter fields to the general relativistic case, we simply make the minimal change

$$\eta_{ab} \to g_{ab} \quad , \quad \partial_a \to \nabla_a$$

in the stress energy momentum tensor. For example, for a perfect fluid it reads

$$T_{ab} = \rho u_a u_b + P(g_{ab} + u_a u_b) ,$$

and for electromagnetism

$$T_{ab} = \frac{1}{4\pi}\left[F_{ac}F_b{}^c - \frac{1}{4}g_{ab}F_{de}F^{de}\right] .$$

A double contraction of The Bianchi identity implies

$$\nabla^a G_{ab} = 0 , . \tag{68}$$

The **Einstein equations in the presence of matter** now read

$$G_{ab} = 8\pi T_{ab} \tag{69}$$

and the matter field equations are again obtained by a generalization of the flat case, $\nabla^a T_{ab} = 0$, which automatically holds because of Eqs. (68,69).

6.5 Problems

1. Show that if a Lorentzian 4-dimensional metric is constant (its coefficients do not depend on the coordinates) in some open set of the spacetime, then it can be brought into the (Minkowski) form,
$$ds^2 = -dt^2 + dx^2 + dy^z + dz^2 \,.$$
(a similar result in fact holds for any signature and manifold dimension).

2. Consider again the metric of a 2-sphere with unit radius,
$$ds^2 = d\theta^2 + \sin^2\theta \,.$$
Compute all the components of the Riemann and Ricci tensors, and the Ricci scalar.

3. Consider cylindrical coordinates $\{\rho, \theta, z\}$, defined in terms of the regular Cartesian ones through
$$x = r\cos\theta \quad y = r\sin\theta \quad z = z \,.$$
Take the flat, standard metric of Newtonian physics, which in Cartesian coordinates is
$$ds^2 = dx^2 + dy^2 + dz^2$$
and explicitly transform it to cylindrical coordinates. Next, restrict the latter to the two-dimensional manifold defined as a cylinder of radius r. Show that even though it appears curved, the Riemann tensor for a cylinder is identically zero. The resolution of this apparent paradox is that its *extrinsic curvature* (which we haven't' defined but essentially describes how a manifold is embedded in a higher dimensional one) is non-zero, but the *intrinsic* geometry of a cylinder is flat.

4. Problem 8, Chapter 3, of Carroll.

5. You are not asked to show it, but any two-dimensional Lorentzian metric can be written in the following form,
$$ds^2 = \Omega^2(t, x)\left(-dt^2 + dx^2\right) \,. \tag{70}$$
We will not discuss this, at least at this point, but the above metric is an example of a so called *conformally flat* one.

   Calculate the components of the Riemann tensor for (70).

6. Prove that every vacuum spacetime ($R_{\mu\nu} = 0$) whose metric has the form
$$ds^2 = -A(x)dt^2 + dx^2 + dy^2 + dz^2, ,$$
where $A(x)$ is an arbitrary positive function of x, is necessarily flat (its Riemann tensor vanishes). Show also that $A(x) = x^2$ is the only solution.

7. Find the Riemann and Ricci tensor components for the two dimensional spacetime
$$ds^2 = -v^2du^2 + dv^2 \,.$$

8. Show, from the conservation of the stress energy momentum tensor, the field equations for a relativistic fluid, Eqs. (62,63).

9. Assuming Eqs. (62,63), derive the non-relativistic version, Eqs. (64,65).

10. Derive the Klein-Gordon equation, (66).

7 Black holes: the Schwarzschild solution

One of the simplest solutions to Einstein's vacuum equations, $G_{ab} = 0$ has remarkable properties and is the topic of this section. Assuming that the metric is static and spherically symmetric, it can be seen that it can be written as

$$ds^2 = -f(r)dt^2 + h(r)dr^2 + r^2 d\Omega^2 ,$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the standard line element of the unit sphere. And, by solving the Einstein vacuum equations, that the functions $f(r)$ and $h(r)$ take the form

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 . \tag{71}$$

where M is an arbitrary constant, called the *Schwarzschild mass*. It might appear odd that we refer to it that way, since the above metric is a solution to the vacuum equations. However, there are several reasons for doing so. One of them is related to a theorem which will not prove,

Theorem 3. *Birkoff's theorem: the Schwarzschild metric is the most general spherically symmetric solution to the Einstein vacuum equations.*

That is, one can drop the assumption of the metric being static, which turns out to be a *consequence* of spherical symmetry in the vacuum case.

Birkhoff's theorem has many consequences, one of them being the fact that the Schwarzschild metric describes the exterior space-time of any spherically symmetric star, some of which we describe in Section 8.1. In particular, we will see that M is the total mass of the star. Another consequence is that there is no gravitational radiation in spherical symmetry: any time dependence of the metric is a coordinate effect.

There are two type of singularities in the Schwarzschild metric as written in Eq. (71): at $r = 0$ and at $r = 2M$. You are asked as part of the homework (Problem 3) to check that there are curvature invariants (scalars) which diverge as $r \to 0$. That is, $r = 0$ is a real, physical singularity. On the other hand, we will explicitly see that the one at $r = 2M$ is a coordinate one, and that by a change of coordinates the metric becomes perfectly regular there. That there is no physical singularity at the so called *Scharzschild radius* $r = 2M$ (the location of the *event horizon*) was not always known. The coordinates (t, r, θ, ϕ) in which the Schwarzschild metric takes the form (71) are (not surprisingly) called *Schwarzschild coordinates*.

We have not defined yet what a black hole is, but the region $r < 2M$ in the Schwarzschild space-time will turn out to be a black hole and its boundary, $r = 2M$, its **event horizon**.

7.1 Geodesics

Recall that geodesics can be obtained from the Lagrangian

$$\mathcal{L} = g_{\mu\nu} u^\mu u^\nu ,$$

where u^μ is the four-velocity (the tangent to the geodesic) and the endpoints in the action are kept fixed. Writing out explicitly such Lagrangian for geodesics in the Schwarzschild metric, considering equatorial motion for simplicity and without loss of generality (since the metric is spherically symmetric, one can always choose the coordinate system to be aligned with the plane of motion of the geodesic),

$$-\kappa = g_{\mu\nu} u^\mu u^\nu = -(1 - 2M/r)\left(\frac{dt}{d\lambda}\right)^2 + (1 - 2M/r)^{-1}\left(\frac{dr}{d\lambda}\right)^2 + r^2\left(\frac{d\phi}{d\lambda}\right)^2 , \tag{72}$$

where $\kappa = 1, 0$ depending on whether the geodesic is timelike ($\kappa = 1$) or null ($\kappa = 0$), and λ denotes any affine parametrization. Since the metric does not depend on t or ϕ (these coordinates are *cyclic*), from the Euler-Lagrange equations,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \dot{t}} &= 0 \Rightarrow \left(1 - \frac{2M}{r}\right) \dot{t} = E \\ \frac{\partial \mathcal{L}}{\partial \dot{\phi}} &= 0 \Rightarrow r^2 \dot{\phi} = L \end{aligned}$$

where E, L are constants of motion and where in order to keep the notation compact we have introduced dots as derivatives with respect to λ. Substituting back into Eq. (73),

$$\frac{1}{2}\dot{r}^2 + \frac{1}{2}\left(1 - \frac{2M}{r}\right)\left(\frac{L^2}{r^2} + \kappa\right) = \frac{1}{2}E^2 \,, \tag{73}$$

which is equivalent to a unit mass particle of energy $E^2/2$ in non-relativistic mechanics with an effective potential

$$V(r) = \frac{1}{2}\kappa - \kappa\frac{M}{r} + \frac{L^2}{2r^2} - \frac{ML^2}{r^3} \,.$$

The second term in the effective potential is as in Newtonian mechanics, the third one is a centrifugal barrier and the last one is new, and dominates over the centrifugal barrier at small r (i.e $V(r) \to -\infty$ for small r).

7.1.1 Timelike geodesics

In this case $\kappa = 1$. Notice that in the weak field regime, i.e. large r, aside from a constant (which does not affect the local motion), the potential is

$$V(r) \approx -\frac{M}{r} \,,$$

which is the Newtonian gravitational potential of a spherical object of mass M. This is one way of interpreting the constant M in the Schwarzschild metric as the total mass of the space-time.

In order to understand the qualitative behavior of geodesics we seek for extrema of the effective potential;

$$0 = \frac{\partial V}{\partial r} = r^{-4}\left[Mr^2 - L^2 r + 3ML^2\right]$$

which has as roots

$$R_{\pm} = \frac{L^2 \pm (L^4 - 12L^2M^2)^{1/2}}{2M} \,.$$

There are two qualitatively different scenarios, depending on the amount of angular momentum L of the test body,

- If $L^2 < 12M^2$ there is no extrema. Notice that for small r

  $$V(r) \sim -\frac{ML^2}{r^3} \to -\infty \text{ as } r \to 0 \,,$$

  while for large r

  $$V(r) \to \frac{1}{2} \,,$$

  i.e. it approaches a positive finite value.

As a consequence, if the particle is initially approaching the black hole or is initially at rest, $\dot{r} \leq 0$, it reaches $r = 2M$ within **finite proper time**. This includes the case of radial ingoing movement ($L = 0, \dot{r} \leq 0$). If the particle is initially outgoing $\dot{r} > 0$, whether it escapes to infinity or bounces back to the black hole depends on its amount of energy. For $E \geq 1/2$ it will escape to infinity and otherwise it will bounce after some radius and reach $r = 2M$ in finite proper time.

Notice that since the metric in Schwarzschild coordinates is singular at $r = 2M$, it is not clear what happens if an observer reaches the event horizon. If there was a physical singularity, the tidal forces would become infinite and his/her life would finish at that point. On the other hand if, as it will turn out to be the case, $r = 2M$ is only a coordinate singularity, we could ignore the metric singularity, notice that the geodesic equation is well behaved there, and conclude that the observer **crosses the event horizon in finite proper time**.

Once the particle is inside the black hole, from the shape of the effective gravitational potential we see that it will reach the physical singularity at $r = 0$ also in finite proper time.

- If $L^2 > 12M^2$ then there are two extrema in the gravitational potential. One can check (do) that R_- corresponds to a local maximum and R_+ to a local minimum. Therefore, there are stable circular orbits ($\dot{r} = 0$) at $r = R_+$ and unstable ones at $r = R_-$. Notice that because

$$R_+ > \frac{L^2}{2M} > \frac{12M^2}{2M} = 6M\,,$$

the so called *innermost stable circular orbit* **(ISCO)** for Schwarzschild is at $r = 6M$. Similarly, one can see that $3M < R_- < 6M$.

The concept of an ISCO is not defined for a generic spacetime, but it is nevertheless useful to consider the following scenario: that one of a particle which is not a test one but instead we qualitatively take into account its gravitational self force as well. If it starts initially on a circular orbit, it will radiate gravitational energy as its orbit shrinks. We can consider an adiabatic approximation of this process, and describe it as a sequence of circular orbits. Once the particles reaches the ISCO, the orbit quickly becomes unstable and the particle falls into the black hole. Numerical simulations of colliding black holes of comparable masses show a similar behavior: once the black holes reach a qualitative "ISCO", the plunge and merger occur in a very short time.

In the Newtonian case all bounded orbits close (not just circular ones), while this is not the case in GR. In the latter, when a timelike stable circular orbit is perturbed around it, it precesses. Quantitatively accounting for the by then well known "anomalous" perihelion precession of Mercury was one of the earliest successes of GR.

7.1.2 Null geodesics

In the case of null geodesics ($\kappa = 0$) the effective potential has always the same qualitative shape, given by

$$V(r) = \frac{L^2}{2r^3}(r - 2M).$$

There is (check) a local maximum at $r = 3M$ and that is the only extremum. Therefore, unstable circular orbits of photons can exist at $3M$.

7.2 A discussion on singularities

Consider the 2-dimensional metric

$$ds^2 = -\frac{1}{t^4}dt^2 + dx^2\,, \quad x \in \mathbb{R}\,, t > 0. \tag{74}$$

The metric has a singularity at $t = 0$, which we will now see is a coordinate one. Through the simple change of variables $t' := 1/t$ it manifestly takes the form of Minkowski (check),

$$ds^2 = -(dt')^2 + dx^2\,.$$

The singularity corresponds to $t' \to \infty$, which cannot be reached by an observer in finite proper time or by a null geodesic in finite affine parametrization. We say that the space-time is **geodesically complete** as $t \to 0$. Furthermore, the original range of coordinates corresponds to $t' > 0$, but since the metric is now regular, we can extend it to $t' \in \mathbb{R}$.

7.3 The Rindler metric

Next consider the metric

$$ds^2 = -x^2dt^2 + dx^2 + dy^2 + dz^2\,, \quad t \in \mathbb{R}\,, x > 0\,. \tag{75}$$

Geodesics terminate with finite proper time at $x = 0$. However, one can see (Problem 5) that the Riemann tensor of this metric identically vanishes; i.e. it is flat and a portion of the Minkowski space-time in disguise. We explicitly show a way of analyzing its global structure, which is interesting by itself but it also leads to a way of analyzing the global structure of the Schwarzschild space-time. The main idea is to use affine parametrization along null geodesics as coordinates. We start with a the tangent vector k^a being null condition,

$$0 = g_{ab}k^ak^b = -x^2\dot{t}^2 + \dot{x}^2$$

where by a dot we refer to a derivative with respect to an affine parameter λ.

7.4 Kruskal coordinates and the global structure of the Schwarzschild space-time

7.5 Problems

1. Check that the Schwarzschild metric, Eq. (71) satisfies the Einstein vacuum equations $G_{ab} = 0$.

2. Problem 3 of Carroll's Chapter 5.9.

3. Compute $R_{abcd}R^{abcd}$ for the Schwarzschild metric and show that it diverges as $r \to 0$, showing that it is a true curvature singularity.

4. Show that the Schwarzschild singularity is spacelike. For this, consider surfaces of constant r with $r < 2M$ and show that at any point on constant r surfaces the vector normal to the surface is timelike. Hint: show that if a surface is given by $f = constant$, then the vector field $\nabla^a f = g^{ab}\nabla_b f$ is normal to the surface.

5. Show that the Riemann tensor of the Rindler metric (75) identically vanishes.

6. Show that any timelike curve inside a Schwarzschild black hole reaches the singularity in finite proper time. Show that such proper time is *maximum* when the observer is free falling (in geodesic motion) – that is, fighting against this doom only makes it worse.

7. Denote by u^a the 4-velocity of an observer in the Rindler metric at constant x, y, z. That is, u^a is the unit tangent to a worldline (the latter parametrized with its proper time so that $u^a u^a = -1$) with constant x, y, z. If the worldline was a geodesic then its 4-acceleration, $a^b := u^a \nabla_a u^b$ would identically vanish. Compute the acceleration of Rindler observers and in particular, its norm, $u^a u_a$.

8. Compute the proper time that it takes for an observer at constant x in metric (74) to reach the $t = 0$ singularity.

9. **[Reading material required]** Two coordinates systems for the Schwarzschild space-time which do not cover the entire extension as Kruskal coordinates do but are simple to obtain and manifestly show that the space-time is regular at the future event horizon are the Eddington Finkelstein (EF) and Painlevee-Gullstrand (PG) ones. Show how to transform from the Schwarzschild metric in Schwarzschild coordinates to EF and PG ones and the final metric expressions in these coordinates and discuss which regions of the Kruskal diagram they cover. One good source to read is `http://arxiv.org/pdf/gr-qc/0001069.pdf`.

8 Static spherically symmetric stars

8.1 Interior metric and matter fields

Here we will consider spherically symmetric stars, where the matter comprising the star is a perfect fluid. Therefore the equations to solve are

$$G_{ab} = 8\pi T_{ab}$$

where

$$T_{ab} = \rho u_a u_b + P(g_{ab} + u_a u_b)\,. \tag{76}$$

Because of the assumption of staticity and spherically symmetry, it can be seen that the metric can be chosen to have the form

$$ds^2 = -f(r)dt^2 + h(r)dr^2 + r^2 d\Omega^2\,,$$

with $d\Omega^2$ the standard metric of the unit-sphere, $d\Omega^2 = d\theta^2 + \sin^2 d\Phi^2$.

As compatibility conditions, in addition to requiring that the pressure P and density ρ depend only on the radius, $P = P(r), \rho = \rho(r)$, we require that the 4-velocity of the fluid, u_a, agrees with the direction of time symmetry, which with normalization included (so that $u^a u_a = -1$) is,

$$u_a = f^{1/2}(dt)_a\,.$$

The relevant components of the Einstein equations that we will need are (primes indicating derivatives with respect to r), which is left as homework (Problem 1):

$$8\pi T_{tt} = 8\pi\rho = G_{tt} = (rh^2)^{-1}h' + r^{-2}(1-h^{-1}) \tag{77}$$

$$8\pi T_{rr} = 8\pi P = G_{rr} = (rfh)^{-1}f' - r^{-2}(1-h^{-1}) \tag{78}$$

$$8\pi T_{\theta\theta} = 8\pi P = G_{\theta\theta} = \frac{1}{2}(fh)^{-1/2}\frac{d}{dr}\left((fh)^{-1/2}f'\right) + \frac{1}{2}(rfh)^{-1}f' - \frac{1}{2}(rh^2)h' \tag{79}$$

Assuming that the density is known, Equation (77) only involves h, and it can be rewritten as

$$\frac{1}{r^2}\frac{d}{dr}\left[r(1-h^{-1})\right] = 8\pi\rho$$

and integrated to give

$$h(r) = \left(1 - \frac{2m(r)}{r}\right)^{-1} \tag{80}$$

where

$$m(r) := 4\pi\int_0^r \rho(\tilde{r})\tilde{r}^2 d\tilde{r} + a \tag{81}$$

with a a constant which has to vanish in order to avoid a conical-type singularity (see Problem 2), $a = 0$.

Next, we solve for h from Eq. (78). Defining Φ through

$$f =: e^{2\Phi}$$

Eq. (78) becomes

$$\frac{d\Phi}{dr} = \frac{m(r) + 4\pi r^3 P}{r\,(r - 2m(r))} \tag{82}$$

which can be solved for if the pressure is known.

Finally, we solve for the pressure assuming its dependence on the density is known. Such dependence is referred to as an *equation of state* (EOS), an example of which is that one of *polytropic* EOS, which refers to $P(\rho) = \kappa\rho^\gamma$, with κ, γ constants. Here we will not explore at all what realistic EOS of states might be (which is an open, difficult question, for highly relativistic systems), but how to solve for the space-time and matter fields once any EOS has been chosen.

It can be seen (Problem 3) that one of the components of the Bianchi identity, $\nabla^a T_{ab} = 0$ is

$$h^{-1/2}\frac{dP}{dr} = -h^{-1/2}(P+\rho)\frac{d\Phi}{dr}\,. \tag{83}$$

Using the solution for h, Eq. (80), and Equation (82) for Φ, it becomes

$$\frac{dP}{dr} = -(P+\rho)\left(\frac{m(r)+4\pi r^3 P}{r\,(r-2m(r))}\right)\,. \tag{84}$$

Having an equation of state, Eq. (84), referred to as the **Tolman-Oppenheimer-Volkoff (TOV)** equation, can be used to solve for the pressure in terms of the density.

Summary of solving for the interior of the star:

1. Specify a density profile, $\rho = \rho(r)$ and an equation of state $P = P(\rho)$.
2. Equations (84) for the pressure is solved for.
3. One of the metric components is given by Eq.(80) and the other one by solving Eq.(82).

Notes:

- The way we have solved for the metric and matter fields is generic, and part of the so called $3+1$ decomposition of Einstein's equations. The bottom-line is that one obtains the relevant equations for the matter fields from $\nabla^a T_{ab} = 0$ and those for the metric from $G_{ab} = 8\pi T_{ab}$. In the case we were able to solve for the matter fields separately [Eq. (84)] and then for the metric. In general that is not the case, though, and the equations are coupled.

8.2 Matching to the outside Schwarzschild metric

We now consider the star as a compact object with the matter fields having compact support. That is, the star has a surface radius R and the matter fields vanish outside it,

$$\rho(r) = 0 = P(r) \quad \text{for } r \geq R\,.$$

Since we are assuming the space-time to be spherically symmetric, from Birkoff's theorem, the space-time outside the star has to be a portion of Schwarzschild. Here we discuss how to match the interior metric as obtained in the previous subsection to the Schwarzschild exterior.

The rr component of the metric is easily obtained,

$$h(r) = \left(1 - \frac{2M}{r}\right)^{-1} \qquad \text{for } r \geq R\,,$$

where, from Eq.(81), we choose

$$M := m(R) = 4\pi\int_0^R \rho(\tilde{r})\tilde{r}^2 d\tilde{r}$$

to match smoothly the interior, Eq.(80).

Once we have fixed the Schwarzschild mass parameter M we need to somehow guarantee that the matching to the outside of the tt component is also smooth. This can be done in two ways. In both cases we first notice that since Eq. (82) is a first order ordinary differential equation, specifying the value of Φ (or, equivalently, f) at any radius determines one and only one solution.

In both cases we start by setting

$$f(r) = \left(1 - \frac{2M}{r}\right) \qquad \text{for } r \geq R\,, \tag{85}$$

1. Solve for $f(r)$ in the interior by solving Eq. (82) "inward" with boundary condition given by

   $$f(r = R) = \left(1 - \frac{2M}{R}\right)\,. \tag{86}$$

   Smoothness between the interior and exterior solution for $f(r)$ is guaranteed.

2. If instead one solves Eq. (82) "outward" by fixing the boundary condition at $r = 0$ (say, by $f(r = 0) = 1$), the solution of such equation at $r = R$ need not satisfy (86). This is not a fundamental obstacle and a rescaling of time inside or outside makes the matching smooth – it is left as homework (Problem 4).

8.3 Physical considerations

Here we point out a number of physical consequences of the physics of a spherically symmetric star as described above.

- In the Newtonian limit, $r^3 P \ll m(r)$ (or, equivalently, $P \ll \rho$) and $m(r) \ll r$, Eq. (82) becomes

  $$\frac{d\Phi}{dr} = \frac{m(r)}{r^2} \tag{87}$$

  which is exactly Poisson's equation for the gravitational potential in Newtonian gravity. Thus, Φ plays such role in the Newtonian limit.

- Similarly, in the Newtonian limit, the TOV equation becomes

  $$\frac{dP}{dr} = -\frac{\rho m(r)}{r^3}\,, \tag{88}$$

  and it is easy to check that if $P \geq 0$, then dP/dr as given above is always smaller than or equal to the relativistic version given by Eq.(84). As a consequence, for any given central density $\rho(r = 0)$, the amount of pressure needed for hydrostatic equilibrium is larger in the general relativistic case than the Newtonian one.

The previous observation becomes more dramatic as shown by the following example and then a general result showing that it is not an artifact: under rather general conditions there is a maximum mass for any fixed radius R of the star. We already have $R > 2M$, the following is a sharper condition.

Example 8. Consider a constant density star:

$$\rho(r) = \rho_0 \qquad \text{for } r \leq R \text{ and zero otherwise.}$$

Then $m(r) = \frac{4}{3}\pi r^3 \rho$ and the Eq. (84) can be integrated to give

$$P(r) = \rho \left[\frac{(1 - 2M/R)^{1/2} - (1 - 2Mr^2/R^3)^{1/2}}{(1 - 2Mr^2/R^3)^{1/2} - 3(1 - 2M/R)^{1/2}}\right] \tag{89}$$

Notice that from the previous equation, the central density $P(r=0)$ needed for equilibrium becomes infinite in the limit $R \to 9M/4$. Therefore we have that in GR, for a spherically symmetric star we have $R > 9M/$. The result is rather general and it is not an artifact of the constant density example, as shown by the following theorem.

Theorem 4. *A spherically symmetric static star of perfect fluid cannot have radius R larger than $9M/$, assuming the following two conditions: $\rho \geq 0$, $d\rho/dr \leq 0$, In particular, no assumption on the equation of state is made. Notice that this a purely relativistic effect, the solution to the Newtonian equilibrium condition, (88) always has finite solutions provided the density is finite.*

Proof:

First, we take a particular linear combination of the Einstein equations,

$$0 = G_{11} - G_{22} = \frac{1}{2}(rfh)^{-1}f' - r^{-2}(1-h^{-1}) + \frac{1}{2}(rh^2)^{-1}h' - \frac{1}{2}(fh)^{-1/2}\frac{d}{dr}\left[(fh)^{-1/2}f'\right] \quad (90)$$

Using the explicit solution for $h(r) = (1 - 2m(r)/r)^{-1}$ and after some algebra (check) Eq.(90) becomes,

$$\frac{d}{dr}\left[r^{-1}h^{-1/2}\frac{df^{1/2}}{dr}\right] = (fh)^{1/2}\frac{d}{dr}\left(\frac{m(r)}{r^3}\right) \quad (91)$$

Since by assumption $d\rho/dr \leq 0$, then the average density $m(r)/r^3$ is also a non-increasing function of radius,

$$\frac{d}{dr}\left(\frac{m(r)}{r^3}\right) \leq 0$$

and from the left hand side of Eq. (91),

$$\frac{d}{dr}\left[r^{-1}h^{-1/2}\frac{df^{1/2}}{dr}\right] \leq 0\,.$$

Therefore, for any $r \leq R$,

$$r^{-1}h^{-1/2}(r)\frac{df^{1/2}}{dr}(r) \geq R^{-1}h^{-1/2}(R)\frac{df^{1/2}}{dr}(R) = \frac{M}{R^3}\,, \quad (92)$$

where in the last equality we have used the fact that the metric at R matches to the Schwarzschild one. Multiplying by $rh^{1/2}$, using the explicit form of $h(r)$ and integrating,

$$f^{1/2}(0) \leq (1 - 2M/R)^{1/2} - \frac{M}{R^3}\int_0^R \left(1 - \frac{2m(r)}{r}\right)^{-1/2} rdr$$

Now, since $d\rho/dr \leq 0$, for the same central density one the mass has to be not smaller than the uniform density case, $m(r) \geq Mr^2/R^3$ and Eq. (92) becomes

$$\begin{aligned} f^{1/2}(0) &\leq (1 - 2M/R)^{1/2} - \frac{M}{R^3}\int_0^R \left(1 - \frac{2Mr^2}{R^3}\right)^{-1/2} rdr \\ &= \frac{3}{2}(1 - 2M/R)^{1/2} - 1/2 \end{aligned}$$

Since by assumption $f^{1/2}(0) > 0$, the last inequality implies

$$M \leq \frac{4R}{9}\,.$$

8.4 Problems

1. Derive equations (77,78,79).

2. The goal of this metric is to illustrate what kind of singularity one might have if the integration constant a in Eq.(81) is not zero. For simplicity and to make the point, consider the standard flat Minkowski metric in cylindrical coordinates,

$$ds^2 = -dt^2 A^2 dr^2 + r^2 d\phi^2 + dz^2 \quad r \geq 0\,, \phi \in [0, 2\pi]\,, z \in \mathcal{R} \tag{93}$$

with the minimal change that A is not necessarily one. The metric has constant coefficients, and is therefore flat everywhere except possibly at the origin, $r = 0$, where the coordinates are singular. We want to show that there is a singularity, and of what type, at the origin unless $A = 1$.

   (a) Transform the metric (93) to Cartesian coordinates and show that it is not differentiable at the origin unless $A = 1$.

   (b) Show that by a rescaling $\phi \to \phi/A$ the space-time (93) describes Minkowski with an an angle "deficit" (i.e. at constant t, z, a "wedge" is removed) when $A^2 > 1$. A similar interpretation holds in the $A^2 < 1$ case.

As a side remark, we point out that the metric (93) with $A^2 > 1$ approximates the exterior space-time to an infinitely long so called *cosmic string*, and it is related to exercise 2 of Carroll's Chapter 5.9.

3. Derive Equation (83).

4. Consider the solution to Eq. (82) with an arbitrary boundary condition at $r = 0$. As discussed below Eq.(86), the solution need not satisfy the matching condition (86). That is, the interior and exterior solutions for $f(r)$ will in general not match at the surface of the star,

$$f_{\texttt{interior}}(r = R) \neq f_{\texttt{exterior}}(r = R)\,.$$

Show that a simple "synchronization of clocks" at the surface of the star, i.e. a rescaling of time either inside or outside the star fixes this in the sense that the condition (86) is satisfied.